DOCTOR OF PHILOSOPHY

An investigation into the implementation of clinical guidelines in the general dental services

Bahrami, Maryam

Award date:
2004

DOCTOR OF PHILOSOPHY

An investigation into the implementation of clinical guidelines in the general dental services

Maryam Bahrami

2004

University of Dundee

An investigation into the implementation of clinical guidelines in the general dental services

Maryam Bahrami

A thesis submitted for the degree of Doctor of Philosophy

University of Dundee
February 2004

I dedicate this thesis to my parents

Contents

Table Index

Page No

Figure and Graph Index

List of Appendices

Acknowledgements

I would like to acknowledge foremost Professor Nigel Pitts for his support throughout the course and provision of a great opportunity for me to undertake this enjoyable experience.

I would especially like to thank Dr Chris Deery and Dr Jan Clarkson, my Supervisors, who provided invaluable support, guidance, help and advice throughout the course.

My special thanks go to Dr Nugent who provided me with her expertise as statistician and offered advice and support as a friend. My appreciation is extended to Dr Colin Tilley for willingly offering his statistical expertise and with all his intuitional and friendly explanations made a regression analysis a simple story. Also Mr Graeme MacLennan who provided his expertise in analysis of the cluster trial and carried out the cluster analysis.

My thanks go also to Mrs Marilyn Laird for her friendly support and help and also for her organisational skills in keeping everybody on time with her colourful charts.

I would like to thank Dr Fay Crawford and all my colleagues in DHSRU for their support and help.

My great appreciation is extended to Miss Hazel Braid for her friendly and super-secretarial assistance in organising this thesis.

Also my thanks go to the general practitioners and their staff who were involved in this study, without their full co-operation this project would not have been possible.

This project was funded by R&D programme. I would like to express my thanks to R&D programme for their financial support. The views expressed here are those of the authors.

I would like particularly to acknowledge with love and gratitude, the unconditional support, devotion and encouragement of my parents and my brothers who have made this experience enjoyable. Finally my special thanks go to my husband who gave me all his love, support, understanding and encouragement. They all never doubted that I would succeed, no matter what obstacles got in the way and they gave me the spirit and the will to continue and I shall forever be grateful to God.

Declaration

I hereby declare that my thesis "An investigation in to the implementation of clinical guideline in the general dental services" is a record of the research I have undertaken at the University of Dundee. I am the sole author of this thesis except where otherwise stated. I have consulted all the references cited.

This research has not been previously submitted for higher degree.

Signed by

Maryam Bahrami

Certificate

We hereby certify that the condition of Ordinance Number 12 and the regulation of the University of Dundee for the degree of Doctor of Philosophy have been fulfilled by this candidate.

Dr C Deery

Supervisor

Dr JE Clarkson

Supervisor

Abstract

Background: An evidence-based approach to clinical practice is advocated to improve the quality of patient care. However there is often a gap between research findings and clinical practice. To address this deficiency there is the need to assist clinicians in accessing and adopting research findings. One possible method of facilitating change in practice is clinical guidelines. It has been shown in medicine that a change in clinical practice in favour of published guidelines is dependent on an active implementation strategy. Consistently effective implementation strategies have not been identified in either medicine or dentistry.

Aim: to investigate the effectiveness of different implementation strategies for evidence based guidelines, using the Scottish Intercollegiate Guidelines Network (SIGN) for appropriate removal of third molar teeth (SIGN 43, 2000).

Design: a randomised-controlled trial employing a 2x2 factorial design linked to multidisciplinary evaluation.

Subjects: 51 volunteer dental practices across Scotland.

Method: Practices were randomly allocated to one of four groups. Pre-intervention data were collected from 49 dental practices. The clinical records of all 16-24 year old patients who attended the practice over a four-month period (August to December 1999) were searched by clinical researchers who were blind to the randomisation. The data extracted included the reason for their attendance and treatment received. This process

was repeated following publication of the SIGN Guideline in April 2000. The post-intervention phase of the project took place between June and October 2000. Data were collected from 46 practices.

Interventions: Mailing of guideline (as control / non-intervention strategy), Audit and feedback (A&F); Computer-aided learning with decision support (CAL-DS), and A&F together with CAL-DS. In addition all practitioners had an opportunity to attend a post-graduate continuation education (PGCE) course on the guideline. Thus the non-intervention/control group mirrored current practice in the dissemination and implementation of the SIGN guideline in primary dental care.

Outcome Measurement: The principle outcome was adherence to the guideline as assessed independently by two researchers. Any disagreement between these evaluators was discussed and an agreement reached.

Results: The overall recruitment rate of practices was 11% of those invited to take part (63 of 565) but this decreased to 8% following the intervention. Prior to the intervention the percent of patients with a problem with their third molar teeth was 7% compared with 22% after intervention. This occurred at the same time as a reduction in the overall number of patients seen by the practices (3342 compared with 1935). A statistically significant reduction in the percentage of patients treated with extraction was detected between the pre- (37%) and post-intervention (27%) phase of this study, ($P=0.02$), where this reduction was not significant for different groups ($P>0.05$).

Compliance with the guideline was 74% of patients pre-intervention and this increased to 78% post-intervention. However, this difference was not statistically significant (P=0.25). The weighted t-test for audit versus no audit (P=0.62) and CAL-DS versus no CAL-DS was not significant (P=0.76). From the multilevel analysis the odds ratio of compliance with the guideline for dentists who experienced audit versus those who did not was 1.28 (95% CI 0.62 to 2.63) and this compares with an odds ratio of 0.84 (95% CI 0.88 to 1.74) for the CAL-DS dentists versus no CAL-DS. For neither was the difference statistically significant. The study was not sufficiently powered to detect an interaction effect so analyses of the main effects only were undertaken. There was however a weak correlation between pre and post cluster level compliance rates (Product Moment Correlation = -0.125, t = 0.81, n = 43, P>0.4). Therefore all analyses were performed on the post intervention compliance rate. All analyses were carried out on an "intention to treat" (ITT) basis.

Conclusion: **There was no statistically significant effect of either CAL-DC or A&F on implementation of these guidelines. This study was unable to show if the CAL-DC and A&F independently had any effect in increasing the general dental practitioners compliance with the guideline but it may have acted as a reinforcement of the guideline messages.**

CHAPTER ONE

Introduction

It has being increasingly recognised that clinical care should be based on the best available evidence. Unfortunately it has also been estimated that only 10%-20% of all clinical practice is based on sound scientific research (The Office of Technology Assessment of the US Congress 1978). So it is not surprising that the investigation and management of a large number of clinical situations are found to be markedly variable and not homogenous (Marriott and Palmer 2000). This variability with clinical care will inevitably result in the provision of suboptimal care and in inefficiency leading to a significant waste of resources. This would be highly undesirable in the context of centrally-based health care provision with limited resources. Part of this problem is a well-recognised delay in the incorporation of research findings into routine practice (Davidoff *et al* 1995, Rosenberg and Donald 1995, Haines and Donald 1998, Bero *et al* 1998, Woolf *et al* 1999). The reasons for this delay are multi-factorial and include information overload, the variability in quality of information which clinical decisions are based on and difficulties in finding the relevant information and in the interpretation of contradictory messages (Rosenberg and Donald 1995, McGlone *et al* 2001).

To confront this difficulty, an evidence-based approach, translating and facilitating the uptake of clinically relevant research findings into practice is being advocated to improve the quality of patient care by reducing harm and maximising benefits (Clarkson *et al* 1999). This Evidence-based practice (EBP) is recognised as a process that restructures the way in which practitioners think about clinical problems. It can provide

answers by ensuring that the best use is made of existing evidence or it can identify an area in which new research is needed (Richard and Lawrence 1995).

However, the integration of this approach into clinical care may require important and substantial changes in professional clinical practice to come in to line with the evidence identified (Richard and Lawrence 1995). In order to achieve these changes in current practice various intervention strategies for the dissemination of research findings have been utilised. One of the advocated methods is the development of evidence-based guidelines in which relevant evidence-based research findings have been filtered in to an accessible format, in an attempt to reduce the observed variability in practice and harmonize clinical activities. Clinical guidelines are commonly defined as "systematically developed statements, which assist in decision making about appropriate health care for specific clinical conditions, services, and outcomes" (Institute for Medicine 1992).

However, development of evidence-based guidelines does not necessarily ensure their use in a daily clinical practice (Feder 1999). Clinical guidelines are not self-implementing and passive dissemination of published guidelines alone is rarely effective in changing the clinical behaviour of practitioners (Effective Health Care 1994, Freemantle *et al* 1999). This is unsurprising given that many factors influence health professionals' behaviour such as the organisational structure, peer group pressure and/or individual variation (Lomas 1991, Haines and Donald 1998, Effective Health Care 1999). There is therefore a need to find the most effective dissemination and implementation strategy or strategies to optimise the incorporation of evidence-based

recommendation into current practice. The dissemination strategies mostly involve raising awareness of existing evidence whereas implementation involves getting the evidence adopted into practice. There are a wide range of potentially effective intervention strategies for the dissemination and implementation of guidelines. These include the use of continuing medical/dental education, opinion leaders, audit and feedback, educational outreach, reminders and multifaceted interventions (Effective Health Care 1999). While these strategies are fundamental to the optimisation of the impact of guidelines in changing clinical behaviour, there is uncertainty about their effectiveness and not enough is known about the best way of transferring the evidence into clinical practice. Much of the evidence to date comes from medical literature and a variable degree of success has been claimed by various intervention strategies but they are not effective under all circumstances (Effective Health Care 1999). It has been recognised that there are "no magic bullets" (Oxman 1995) and the choice of appropriate strategies should be considered in line with the potential barriers and facilitators that may influence the desired changes in a clinical practice (Effective Health Care 1999).

Very few studies have investigated this process in dentistry (O'Brien *et al* 2000, Goodey *et al* 2000, Kay *et al* 2001). Although many factors may be similar between dental and medical practice, there may also be considerable differences, given the different organisations and funding structures (McGlone *et al* 2001). Also, the effect of these interventions in medicine may not be directly applicable to the experience of dental profession. Hence, there is a need for research into dentistry while using the experiences from the medical profession to evaluate the effectiveness of different intervention

strategies to change the dental practice. The need for the development of effective strategies in dental practice has been acknowledged in line with the concept of "clinical governance" introduced by the UK government to improve the quality of clinical care throughout the NHS (Secretary of State for Health 1997, Department of Health 1998).

There was scope for third molar teeth to be identified as an appropriate subject in dentistry for guideline development for a number of reasons. There appeared to be a substantial variation in the management of patients with a third molar problem (Gilthorpe *et al* 1997, Landes 1998, Eklund 2000). The removal of third molar teeth is frequently associated with considerable morbidity (Ogden *et al* 1998, Song *et al* 1997), and is a common surgical procedure in the UK so is therefore associated with significant expenditure (Effectiveness Matters 1998).

Several guidelines have been developed to address the variability in practice with respect to the management of third molar teeth, examples of which are the National Institutes of Health (NIH) Consensus Criteria (1980), the documents produced by a working party of the Faculty of Dental Surgery of Royal College of Surgeons of England (1997) and the National Institute for Clinical Excellence guidelines (NICE) (1999). The most recently published guideline related to third molar teeth was Scottish Intercollegiate Guidelines Network (SIGN) guideline "for unerupted and impacted third molar teeth" (SIGN 43, 2000).

The proposed study was aimed at examining the dissemination and implementation of this guideline in primary dental care. The study took the form of a cluster randomised

controlled trial employing a 2x2 factorial design in Scottish primary dental care settings to compare and evaluate various dissemination and implementation methods for the guidelines. Employing SIGN 43 Guideline (SIGN 43, 2000) on the management of unerupted and impacted third molar teeth as a model.

The strategies being assessed were the mailing of the SIGN Guideline (as control/non-intervention strategy); audit and feedback (A&F) and computer aided learning with decision support (CAL&DS) (as strategies under investigation). In addition all participating practitioners had an opportunity to attend a postgraduate continuing education course (PGEC) on the guideline.

The aim of the study was to evaluate the effectiveness of different dissemination and implementation strategies for an evidence-based guideline for the appropriate management of third molars.

Additional economic and behavioural evaluations of the respective implementation strategies have been completed by other research teams in the Dental Health Services Research Unit (DHSRU) in the University of Dundee and the Psychology Department in the University of St Andrews. Their work will not be covered in this thesis but their outcomes can be found in the executive summary report of the study to the NHS Research and Development funding body (R&D Report, R2-64, 2002) and also in a paper by Bonetti *et al* (2003) in Appendix I.

In a comparison of the effectiveness of strategies under investigation for dissemination and implementation of the guideline (SIGN 43, 2000), our Null Hypothesis assumed no difference between the intervention and non-intervention/control groups.

In order to achieve the aims of this study two main research questions were addressed.

1. What is the overall compliance of the participating practitioners with the SIGN Guideline (SIGN 43, 2000) at both pre and post intervention?

2. Are the interventions under investigation (*i.e.* audit and feedback or computer aided learning programme (CAL)) effective in changing the practice of the participants with respect to the management of third molar teeth?

The following research questions were also considered:

3. Whether the current practice of mailing guidelines and postgraduate courses (non-intervention group/control) would significantly influence the compliance of the recruited general dental practitioners (EBP) with the guideline?

4. Whether the interventions under investigation would have more impact than the current practice of mailing guidelines and postgraduate courses (non-intervention group/control) in altering evidence-based practice (EBP) for third molar management?

5. If any of the participating practices withdrew from the study, whether there was any difference between the dentists who withdrew from the study and those who continued?

6. What are the most frequent third molar pathologies diagnosed for the study patients (aged 16-24 years old) in both phases of the study?

7. Concurrent with study, whether there was any difference between the sample study population and other dentists in Scotland and whether there was any change in the rate of third molar removal in Scottish general dental practices during the experimental period?

8. Whether the participating GDPs complied with specific parts of the guideline on medical history or radiographical examination?

This thesis consists of five main chapters. Chapter 2 reviews the current literature and is presented in four sections. The first section (2.1) intends to present the principles of evidence-based practice and an understanding of the steps involved in this process. This section also discusses methods of locating the evidence and provides an overview of the organisations which systematically prepare, maintain and disseminate the evidence. Section 2 of Chapter 2 (2.2) of the literature review will continue to discuss the pivotal role of evidence-based clinical guidelines, the process of their development, their classification schemes and their advantages and disadvantages which demonstrates that guidelines are one of the most encouraging but complex dissemination strategies in

evidence-based practice. Section 2 (2.2) will further provide an account of various implementation effectiveness studies, in order to give an insight into the different intervention strategies and their effectiveness. There is a great deal of information on behavioural theories and potential barriers and facilitators for the implementation of evidence (Haines and Donald 1998), which is beyond the scope of this literature review; nevertheless a brief reference to this is made.

Section 3 of Chapter 2 (2.3) gives a brief description of a number of potential study designs which are commonly used for the evaluation of implementation strategies. The occurrence of systematic methodological errors and bias can jeopardise the validity of any study, so this section (2.3) discusses the subject of validity and contains a summary of the most common biases encountered in research of this nature. Inadequate reporting of randomised controlled trials (RCT) can lead to errors when interpreting their results. The Consolidated Standards of Reporting Trials (CONSORT) recommendations were published in 1995 to assist with the improvement of the quality of RCT reports. An overview of the CONSORT report is also outlined in this section (2.3).

As this study uses the SIGN Guideline (SIGN 43, 2000), Chapter 2, Section 4 (2.4) gives an account of the epidemiology, aetiology, indications and complications of the removal of third molar teeth. The aim is to provide the reader with background information about third molar teeth. Also an account of the variation in the provision of care for patients with a third molar problem is given.

Clinical decision-making is a complex process and the discrepancy in decision-making among clinicians could originate from two main sources: variability in the perceptual or judgemental decision-making processes. In a continuation with an account of the variation in the provision of dental care, Section 4 (2.4) carries on to describe these processes briefly.

This leads on to Chapter 3, which describes the scientific methods employed in this trial. The penultimate chapter, Chapter 4 presents the results of the analysis of the data in five sections. In the first section (4.1), the results of the descriptive analysis of the data are presented. This is followed by section 2 (4.2), where the results of the data analyses examining the main effect of the interventions under investigation are provided. The result of analysis is presented at "practice" and "patient" level. The third section of Chapter 4 (4.3) contains the sub-analysis at "patient" level, examining the changes within the groups following the intervention phase. The fourth section (4.4) outlines further results of the descriptive analysis of patients' characteristics. This is followed by Section 5 (4.5) where the data on the rate of third molar removal by Scottish dentists are presented. Finally a brief summary of the results is presented in Section 6 of this Chapter (4.6).

The final chapter, Chapter 5 discusses the method, results and implications of this study, with recommendations made for future investigations.

In conclusion, one of the key components of the UK National Health Service policy over recent years has been the improvement of quality by reducing the existing variation

in primary care practice by the active dissemination of relevant evidence (Scally *et al* 1998). In dentistry less attention has been given to implementation of research findings and therefore the effectiveness of potential intervention strategies in primary dental care is unknown (McGlone *et al* 2001). Hence, there is the need to investigate the effectiveness of potential dissemination and implementation strategies in dental primary care. This will be of value to the profession and health care organisers alike, but most importantly to the patients.

CHAPTER TWO

Literature Reviews

Section 1: Evidence-Based Practice

2.1 Introduction

Clinicians strive to provide their patients with optimal care based on available research. Nevertheless, knowing whether these research findings are valid and appropriate is often hampered by uncertainty. Assisting clinicians to find the best, unbiased and up-to-date evidence is the object of evidence-based practice.

To achieve this objective however, there is a need for a systematic review of the available research which involves locating, assessing, and preparing the evidence. This process is resource-intensive and time consuming but there are organisations which make efforts, nationally and internationally, to minimise the unnecessary duplication of reviews and present them in a concise, clear, accessible and efficient format to the practising clinician.

This section of the literature review seeks to provide an overview of evidence-based practice and to describe the process of practising evidence-based care in daily clinical situations. It then goes on to describe those national and international organisations that assist in achieving the goals of evidence-based practice by preparing the systematic reviews.

2.1.1 Why Evidence-Based Practice

For decades, there has been a growing awareness of a gap between what ought to be done as clinicians and what is actually done in practice which translates into the differences between the findings of research and clinical practice (Davidoff *et al* 1995, Rosenberg and Donald 1995, Haines and Donald 1998, Bero *et al* 1998, Woolf *et al* 1999). The two driving forces underlying this gap are – one, the quality of the evidence which clinical decisions are based on and two, the sheer volume of information.

Research has shown that many current clinical interventions have been based largely on intelligent guesswork, dated primary training and individual clinical experience and skills (Rosenberg and Donald 1995). The office of the Technology Assessment of the US Congress (1978) stated that only 10 - 20% of all procedures used in medical practice were based on scientifically and statistically sound research. Many years later, in United Kingdom patient-based studies, Ellis *et al* (1995) and Gill *et al* (1996) demonstrated that approximately 80% of clinical decisions in medicine were supported by evidence. They did, however, suggest that similar studies should be conducted in other specialities.

Ellis *et al* in 1995, conducted a review of treatments given to 109 consecutive patients over a short period in a single general medical ward and searched the medical sources for randomised controlled trial (RCT) evidence evaluating the effectiveness of the interventions. They found that 53% of the treatments were supported by the findings of RCTs and 29% by non-experimental evidence. Overall 82% of the patient management interventions that they studied were based on high quality scientific evidence (Ellis *et al*

1995). Similarly in a retrospective review of consultation records for 101 consecutive patients in 1996, in one suburban general medical practice, Gill and his colleagues demonstrated that 30% of the interventions were based on randomised controlled trial evidence and 51% on convincing non-experimental evidence. They concluded that 81% of interventions were based on evidence (Gill *et al* 1996).

Therefore it is safe to assert that at least a proportion of clinical decisions are not based on sufficient scientific evidence (Bouchier 1997).

Secondly, keeping up-to-date with literature in daily clinical practice is becoming increasingly more difficult for clinicians (Marriott and Palmer 2000). The problem is the volume of literature, since the clinician who tries to keep up with relevant journals faces the task of examining 17 articles daily every day of the year (Haynes 1993). In addition, most of the articles are published in inaccessible places, in different languages or are not even published at all (i.e. reports, working papers, dissertations and conference abstracts which often have very limited dissemination and/or are not included in bibliographical retrieval systems and they are termed "grey literature") (Chalmers *et al* 1992, Last 1995).

The literature, published or unpublished can be seriously biased, poor quality, or irrelevant (Chalmers *et al* 1992, Davidoff *et al* 1995, McAuley *et al* 2000). For instance, over 95% of articles in medical journals failed to reach the minimum standards of quality and clinical relevance (Haynes 1993).

Research with statistically significant results is more likely to be submitted and published than work with neutral or negative findings (Clarkson *et al* 1999). Smart, in 1964 demonstrated this by randomly selecting 169 abstracts presented at the American Psychology Association annual meetings and PhD dissertation abstracts in Psychology. He found those with positive results (75%) were more likely to be published than those with negative results (Smart 1964). It has been reported by McAuley and colleagues that in a review of a random sample of 135 meta-analyses only 33 publications included both grey literature and primary studies. They reported that the exclusion of grey literature from meta-analyses may result in an overestimate of an intervention effect by an average of 12% (McAuley *et al* 2000). This can have implications for clinician practice and ultimately for patient care.

Furthermore, some well-designed studies on relevant subjects are analysed and presented in ways which are hard to implement into the daily clinical practice, or they fail to address the problems which arise in real situations (Godlee 1998). Hence, health professionals are confronted with the difficult task of finding relevant information for their daily clinical problems and at the same time coping with a rapidly changing body of relevant information (Rosenberg and Donald 1995) as well as difficulty in interpreting the contradictory messages (McGlone *et al* 2001).

So, how can the gap between "clinical practice" and "good clinical research" be addressed and how can clinicians gain easy access to high quality, reliable summaries of research findings?

Evidence-based practice (EBP) attempts to fill this gap by the process of systematically finding, appraising, and using current evidence as the basis for clinical decisions, by making research findings more accessible to clinicians, and more clearly applicable to daily practices (Bader *et al* 1999).

2.1.2 Evidence-based paradigm

Evidence-based practice is ''the conscientious, explicit, and judicious use of the current best evidence in making decisions about the care of individual patients'' (Sackett *et al* 1996). Its philosophical origin extends back to the mid-19th century in Paris when Pierre-Charles-Alexander Louis used statistics to measure the effectiveness of bloodletting, the results of which helped to put an end to the practice of leeching (Rangachari 1997). In the UK, one of the most important advocates of EBP was Archie Cochrane. His experiences in the prisoner of war camps, where he conducted trials in the use of yeast supplements to treat nutritional oedema, influenced his belief in a reliable and scientifically proven treatment. In 1972, he published his book "Effectiveness and Efficiency". Cochrane advocated the use of randomised control trials (RCT) (RCT will be discussed in detail later in this chapter section 2.3.1.4) as the gold standard in the research of all medical treatment, where possible, and the systematic reviews of these trials to gather evidence-based information. An interest in EBP has flourished in the last few decades of the 20th century.

Evidence-based practice involves integration of the best available clinical evidence derived from systematic research with clinical expertise, and the patients' values and expectations but it is not a "cookbook" approach to practice. It informs, but never

replaces, clinical judgement (Sackett *et al* 1996). The best available clinical evidence is clinically relevant research, which is focused on patient well-being and evaluates the precision of diagnostic tests and prognostic indicators and the effectiveness and safety of therapeutic, preventive, and rehabilitative procedures (Sackett 1997). EBP also has a role in assessing the cost-effectiveness of these interventions. However, it is not a substitute for rationing and often results in practice that, despite being more cost-effective, has greater overall costs (Hunter 1996).

Evidence-based health care (EBH) is a newer term, which broadens the evidence-based practice techniques to other aspects of health care delivery.

More generally, evidence-based practice is a global movement in all the health care disciplines and the application of its principles is being seen increasingly across the different fields of health care such as medicine, dentistry, nursing, pharmacy and other professions allied to medicine.

2.1.3 The process of applying Evidence-Based Care (EBC) in clinical practice:

Practising evidence-based care is a process of lifelong self-directed learning which involves making logical decisions in diagnosis, prognosis and treatment, and is supported by the current best research findings and relies on the clinical skill and experience of individual clinicians on patient care (Bader *et al* 1999). EBP asks questions, finds and appraises the relevant data, and captures that information for everyday clinical practice. It is not an ivory tower endeavour for academics. Rather, it

is the domain of practitioners. The evidence-based skills can be easily learned by clinicians of varying backgrounds and at any stage in their career to guide their practice (Rosenberg and Donald 1995). Furthermore, it is not based on old research findings. Rather, it is based on sound and updated scientific research evidence, which makes for sound daily clinical practice (Rosenberg and Donald 1995). Five steps have been identified in the process of applying EBC that should become a part of day to day practice (Socket *et al* 1996):

- Converting the problem into a concise question that addresses uncertainties in patient management
- Searching the literature to identify the highest-quality relevant clinical evidence
- Evaluating and critically appraising the evidence for its validity and clinical applicability (Hunt *et al* 2000)
- Implementing the result in clinical practice
- Finally, evaluating the clinicians' performance (Rosenberg and Donald 1995)

These steps are considered in detail below.

2.1.3.1 Define a Clinical Question

The first skill in EBP is to identify uncertainties in patient care and formulate these into focused clinical questions. A focused clinical question is one which clearly addresses a patient or a problem, and considers the intervention and the clinical outcome of interest (risks/benefit and cost effectiveness) (Sackett *et al* 1997). The questions can cover a wide spectrum of issues (i.e. clinical findings, aetiology, differential diagnosis, diagnostic tests, prognosis, treatment, prevention, outcome and self-improvement,

quality of care, or even health economics (Sackett *et al* 1997, Rosenberg and Donald 1995). However, they should always focus on patient well-being. The questions ought to be balanced against a number of factors such as the feasibility of finding an answer applicable to the relevant interests of the clinician, and whether the question is likely to be raised by subsequent patients and could therefore be of value to them.

2.1.3.2 Searching for the best evidence

Once a question has been formulated, the next step in undertaking EBP is the identification of all the best available evidence. Relevant sources of information must be sought to identify all the available literature that will help in answering the question. Busy health professionals need timely, effective and easy access to valid, relevant and current research findings when they are making decisions. There are different strategies and different types of resources involved in answering a clinical question.

The first line of information for most clinicians is the use of textbooks, a traditional source of synthesised information. They are easy to use and suitable for answering general questions. Unfortunately they are often out-of-date even at the time of publication, can be opinion based, and often lack a systematic approach to evidence, summarisation and citation (McKibbon *et al* 1996). However there are textbooks available in electronic formats, for example in medicine electronic textbooks such as UpToDate, Scientific American Medicine and eMedicine that can provide general background information. These are well referenced and updated frequently (Hunt *et al* 2000).

Another resource is a literature search, either in paper format or on an electronic database which can provide background information on many topics. Nevertheless, the range of literature available and the complexity of such databases make searching somewhat difficult and time consuming. It has been proven that a traditional literature search involves intractable indexing problems that undermine the success of searches for the relevant studies (low sensitivity), while retrieving many studies that are not relevant (low specificity) (Haynes *et al* 1997). Widespread global access to electronic databases means comparatively rapid identification of relevant medical literature as opposed to the paper format. An example of such a database is MEDLINE. The computerised bibliographic database MEDLINE, maintained by US National Library of Medicine, is a comprehensive collection of articles that is constantly being updated and includes more than nine million citations of both clinical and pre-clinical studies. Ready accessibility makes it an attractive database for finding health care information. MEDLINE's disadvantages are its size and the range of publications it encompasses. Searching MEDLINE needs careful thought and a thorough knowledge of how the database is structured and how its publications are indexed (Hunt *et al* 2000). The OMNI project (Organising Medical Network Information) is another useful Internet-based resource that provides the UK and worldwide coverage of resources in medicine, biosciences and health management. The lead body is the National Institute for Medical Research Library, with contributions from the medical libraries of Nottingham University, Cambridge University, the Royal Free Hospital School of Medicine, the King Edward Hospital Fund and the Wellcome Centre for Medical Science (http://www.omni.ac.uk).

Alternatively, strategies such as "systematic reviews" have been employed. These allow clinicians and others to digest clinically important findings from a vast amount of medical literature (Mulrow 1994). Systematic reviews are ''systematic scientific summaries of available scientific evidence that use quantitative methods to analyse and summarise the results that address targeted clinical questions'' (Woolf 1996). Like original research projects, systematic reviews start with a hypothesis that is answered through collection, appraisal and analysis of published and unpublished research (grey literature). The evidence is prepared methodically with a comprehensive systematic assessment. It differs from the inconsistent and haphazard traditional reviews based on individual opinions which can be prone to error and bias (Mulrow 1987). It is essential that any systematic review be free of systematic errors (bias) and random errors (chance). (Bias will be discussed in detail later in this chapter in section 2.3.4).

Systematic reviews are conducted by a collaborative multidisciplinary team in the appropriate clinical disciplines and by methodologists trained in searching, appraising and summarising all evidence, whether published or unpublished (Bader *et al* 1999). Reviews, which have undergone these critical appraisals during their development and thus contain pre-filtered evidence-based information, are available to help answer clinical questions (Mulrow 1994). They also highlight areas of insufficient evidence. In response to demand for easy access to these pre-filtered sources of evidence, centres were established as part of the information system strategy. The Cochrane Collaboration was established in the UK in 1992 and its international collaboration commenced in 1993 (detailed later in this chapter in 2.1.4.3c). The Cochrane library attempts to make this type of information accessible and focuses primarily on systematic reviews of

controlled trials, or RCTs, of therapeutic interventions. Well-conducted RCTs are considered the most valid research methodology and are the gold standard in systematic reviews of therapeutic interventions. In its simplest form RCT means that subjects are randomly assigned to one of two groups: intervention or non-intervention, and the outcomes are then compared (RCT will be discussed in detail later in this chapter, in section 2.3.1.4).

Other establishments with similar functions have been set up. For instance, the Agency for Health Care Research and Quality (AHCRQ) is an agency which was set up in 1989 in the United States. In 1997 it launched its initiative to promote evidence-based health care. It has funded 12 Evidence-based Practice Centres (EPCs) in the United States and Canada, which reported their first set of reviews in 1999 (AHCRQ web page: http://www.achcpr.gov.2000). In the UK, the National Health Service Centre for Reviews and Dissemination (CRD), established in 1994, is another source of high quality systematic reviews of health care interventions (detailed later in Section 2.1.4.2).

Although these initiatives fulfil an essential role by producing regularly up-dated systematic reviews of evidence of clinical effectiveness, an inevitable delay exists between the publication of important evidence and its incorporation into review. Secondary publications fill this gap. There are rapidly expanding sets of journals that provide access to selected, appraised and combined results from primary information sources. Two of which are "the American College of Physicians' ACP Journal Club" and the "Evidence-Based Medicine" journals, which were launched respectively in 1991 and 1995. In January 2000 these two journals were combined into one journal - the

"ACP Journal Club" - in North and South America but "Evidence-based medicine". the UK version of the journal, continues to be available outside America (Hunt *et al* 2000). The aim of these journals is to locate and summarise the small number of clinical articles that are methodologically sound and clinically relevant and publish them as structured abstracts along with an accompanying commentary provided by experienced clinicians who put the study's findings into a clinical perspective. Every five years each article is reviewed to make sure it has not become outdated in light of new and more recent evidence. ''Best Evidence'' is the electronic version of these two paper-based abstract journals. (Best Evidence 2000 Issue 4 covers all issues in the ACP Journal Club from 1991 to 1999 as well as all issues of Evidence-based Medicine from 1995 to 1999. "Best Evidence" is also available in CD-ROM format.) It was first launched in 1998, all its materials are updated in a five-year cycle and it is very easy to search. However its coverage is limited.

For different health care specialities, only a limited number of summaries in these two journals may be relevant. So several other journals, which are dedicated to various sub-specialities, have also been published. Many more of these new types of journals are being developed so that eventually most clinical specialities will have their own journal. Evidence-Based Dentistry is one of these journals and was launched in November 1998. Its purpose is to alert dental clinicians to important advances in dentistry by selecting evidences from dental literature which are relevant, valid and reliable (Lawrence 1998).

2.1.3.3 Critically appraising the evidence

Once potentially useful evidence has been found it must be critically appraised to establish its validity, quality and relevance to the clinical question (Rosenberg and Donald 1995, Straus and Sackett 1998). Consequently, a structured but simple method, the "critical appraisal", has been developed by several teams working in North America and the United Kingdom. This is a process to help clinicians to assess accurately all the available evidence in terms of its validity and clinical applicability. Mastering this method involves learning how to ask a few key but simple questions to establish the quality and relevance of findings to a particular clinical problem (Rosenberg and Donald 1995). For example, questions about diagnostic tests (was there an independent, blind comparison with a gold standard of diagnosis?); about treatment (was the allocation of patients to treatments really randomised?); about prognostic markers (was an appropriate sample of patients gathered at a uniform point in their illness?) and clinical guidelines or other strategies for improving the quality of care (Sackett *et al* 1997, Oxman *et al* 1993).

Furthermore, to evaluate the evidence and assess its validity, grades have been developed depending on the strength of the evidence and on the effectiveness of a treatment. These grades account for the type and quality of the study design and the variability of study results (Liddle *et al* 1996). Thus a systematic review of randomised control trials that shows consistent results would be graded as providing higher quality evidence than a review of randomised control trials that presented variable results without a good explanation of the variability (Sheldon *et al* 1998).

The next important step in the process of incorporating the EBP into the daily management of patients is the implementation of the evidence.

2.1.3.4 Implementing the evidence in clinical practice

Having identified evidence that is valid and relevant, the results of critical appraisals can then be applied in practice. In other words the clinical message is extracted from the critically appraised evidence and applied to the individual patient. This involves integrating the evaluated evidence with the expertise of the individual clinician who will take into account the individual patient's predicament, rights and preferences in making his clinical decision. There is a range of interventions available for implementing the evidence which will be considered in more detail in Section 2.2.2.1 of this chapter.

2.1.3.5 Evaluation of performance

Having introduced the evidence to their clinical practice, clinicians can then evaluate their decisions and performance to see if the clinical outcome has improved. They can evaluate their progress at each stage by asking if evidence was effectively appraised and if the integration of the evidence with their clinical expertise and individuality of their patient helped them with a rational and acceptable management strategy. This self-evaluation allows clinicians to focus on the areas that may need improvement in future (Straus and Sackett 1998).

Having described the different stages in the process of undertaking EBP, it is important to keep in mind that new evidence is constantly evolving and may contradict older findings. Therefore the process of evidence-based practice is based on a continuous

review of research findings. New evidence is continuously appraised and made available to busy clinical practitioners, policy makers and planners. Consequently different organisations nationally and internationally such as the Cochrane Collaboration and the Agency for Health Care Research and Quality (AHCRQ) are working to collate databases of good quality reviews and helping to avoid unnecessary duplications of these reviews to promote evidenced-based practice and the delivery of effective health services.

2.1.4 The Research and Development Strategy of the UK National Health Service

The systematic review of the evidence using well-established and rigorous strategies assist clinicians and policy makers in knowing whether the conclusions of a review of the evidence are valid and whether the recommendations to practice guidelines are sound.

Nevertheless, undertaking these systematic reviews are resource-intensive and time-consuming. Therefore an attempt has been made nationally and internationally to minimise the duplication of the reviews. Several organisations have been involved in undertaking this type of review and presenting their results widely such as the Research and Development strategy of the National Health Service and the Cochrane Collaboration.

The Research and Development Strategy of the National Health Service seeks to promote a knowledge-based health service (Bouchier 1997) and the Research and

Development Information Systems Strategies consist of three components (Sheldon and Chalmers 1994):

The National Research Register

The NHS Centre for Reviews and Dissemination (CRD)

The Cochrane Database of Systematic Reviews

2.1.4.1 The National Research Register

The National Research Register gathers information on funded and current research for Health Services in the United Kingdom. This database will enable planners, providers and fund-holders to determine whether their research information requirements are being addressed and should allow them to make direct contact with the relevant research workers (Sheldon and Chalmers 1994).

2.1.4.2 The NHS Centre for Reviews and Dissemination (CRD)

The NHS Centre for Reviews and Dissemination is based in the University of York and funded by the NHS Executive and Health Departments of Scotland, Wales and Northern Ireland. It was established in January 1994 to provide the NHS with important information on the effectiveness and cost effectiveness of health care intervention as well as the delivery and organisation of health care. CRD helps to promote evidenced-based clinical practice in the NHS by offering rigorous and systematic reviews on selected topics which has resulted in a database of good quality reviews.

The UK Cochrane Centre and CRD are part of the single Information Systems Strategy supporting the NHS R&D programme. The CRD plays an important role in assisting the Cochrane Collaboration in disseminating the contents of The Cochrane Database of systematic Reviews to the NHS (Sheldon and Chalmers 1994).

CRD collaborates with a number of health research and information organisations across the world and is the UK member of the International Network of Agencies for Health Technology Assessment (INAHTA). It produces a database of HTA projects and publications. This network has been established to promote an exchange of information and to improve the quality of reviews of health technologies and the dissemination of these assessments (Sheldon and Chalmers 1994).

2.1.4.3 The Cochrane Collaboration

The background to the origin of the Cochrane Collaboration is the recognition of an unmanageable amount of information, which overwhelmed health care professionals, researchers, policy makers and consumers alike. As has been briefly already mentioned, in 1972 Archie Cochrane the British epidemiologist who inspired this collaboration, criticised the medical profession for its ignorance about the effects of health care. He recognised that there were no reliable reviews of available evidence for people to make informed decisions about their health care (Cochrane 1972).

Cochrane's suggestion that the methods used to prepare and maintain the reviews should be applied more widely was taken up at an international level. Subsequently, in 1992, as part of the British National Health Service (NHS) Research and Development

(R&D) Programme, ''The UK Cochrane Centre''(UKCC) was established (Sheldon and Chalmers 1994). Later on in 1993, ''The International Cochrane Collaboration'' was co-founded by 11 countries. The UK Cochrane Centre based in Oxford is one of 15 such centres around the world which provide the infrastructure for co-ordinating the Cochrane Collaboration. It facilitates and co-ordinates the preparation and maintenance of systematic reviews of randomised control trials in health care. Cochrane reviews (the principal output of the Collaboration) are published in the Cochrane Library which is an electronic journal available on disk and CD-ROM. It is an essential resource for organisations and individuals wishing to improve the quality of medical practice by using the accumulated evidence published in the primary medical literature. This database is regularly updated by Cochrane Collaboration Review Groups (Antezak-Bouckoms and Shaw 1994). The Cochrane Library consists of different components which include: The Cochrane Database of Systematic Reviews (CDSR), The Database of Abstracts of Reviews of Effectiveness (DARE) and The Cochrane Controlled Trials Register (CCTR)

2.1.4.3a The Cochrane Database of Systematic Reviews (CDSR)

This is a rapidly growing collection of systematic reviews of medical literature, in particular randomised control trials. Evidence is included or excluded on the basis of explicit quality criteria in order to minimise bias. In many cases the data is combined statistically into a meta-analysis. This is a statistical technique summarising the results of several studies into a single estimate, giving more weight to results from larger studies and increasing their statistical power. Meta-analyses assist by providing analysis compiled from numerous studies, some of which may be too small to produce reliable

results individually.

2.1.4.3b The Database of Abstracts of Reviews of Effectiveness (DARE)

This is a collection of structured literature abstracts of systematic reviews from around the world which have been critically appraised by reviewers at the British National Health Service (NHS) Centre for Reviews and Dissemination at the University of York. DARE also includes records of abstracts of reports of health technology agencies worldwide, and abstracts of reviews in the ACP journal Club and Evidence-Based Medicine journal (Sheldon and Chalmers 1994).

2.1.4.3c The Cochrane Controlled Trials Register (CCTR)

This is a bibliographic database of controlled trials identified by contributors to the Cochrane Collaboration as part of an international effort to hand-search the world's journals and create an unbiased source of data for systematic reviews. The register includes references to reports in conference proceedings not indexed in other literature databases.

2.1.4.4 The aims and objectives of the Cochrane Collaboration

This international organisation aims at helping people to make well informed decisions about health care by preparing, maintaining and ensuring the accessibility of systematic reviews of the benefits and risks of health care interventions (Alderson 1998). The principal goal of the Cochrane Collaboration is to ensure that high quality, rigorous and up-to-date systematic reviews are available across a broad range of health care disciplines. The eight basic principles on which the Collaboration is built are (Alderson

1998):

1. Collaboration
2. Building on the enthusiasm of individuals
3. Avoiding duplication
4. Minimising bias
5. Keeping up-to-date
6. Ensuring relevance and access
7. Continually improving the quality of its work
8. Continuity

Over 7000 Cochrane Collaboration members from all over the world are organised into 50 specialised review groups covering each area within health care. The Cochrane Effective Practice and Organisation of Care Group (EPOC) and The Cochrane Oral Health Review Group are two examples of the Collaborative Review Groups (CRG). These two examples have been chosen from the 50 collaborative review groups as they relate to our proposed study and dentistry. EPOC aims to help people make well-informed decisions about health by providing systematic reviews of the evidence of effectiveness of different implementation strategies. The Cochrane Oral Health Review Group is one of the landmarks in the history of evidence-based practice in dentistry and is engaged in conducting systematic reviews in dentistry, focusing on Evidence-Based Dentistry.

2.1.4.5 Cochrane Effective Practice and Organisation of Care Group (EPOC)

EPOC is a Collaborative Review Group of the Cochrane Collaboration. This is an international organisation concerned with preparing, maintaining and ensuring accessibility to systematic reviews of the effectiveness of interventions designed to improve professional practice and the delivery of effective health services. This includes various forms of continuing education, quality assurance, informatics, financial, organisational and regulatory interventions that can affect the ability of health care professionals to deliver services more effectively and efficiently through these systematic reviews (Oxman 1999).

2.1.4.6 The Cochrane Oral Health Review Group

The Cochrane Oral Health Review Group was set up in 1994 and is one of the speciality-based review groups within the Cochrane Collaboration. It is based in Manchester, UK and has a broad range of contributors such as dentists, patients, statisticians, and epidemiologists. The goal of this group is to prepare and maintain systematic reviews of RCTs related to oral health. The scope of this group covers the prevention, diagnosis, treatment, and rehabilitation of oral, dental and craniofacial diseases and anomalies. The Cochrane Oral Health Group has several registered protocols and has completed reviews in many areas such as orthodontic and oral surgery, oral medicine and periodontology (*e.g.* Orthodontic Treatment for Posterior Crossbites and Interventions for Treating Oral Lichen Planus). These reviews are being updated and added to all the time. They have also collected more than 11,000 citations to clinical trials in dentistry (Antezak-Bouckoms and Shaw 1994).

Other organisations which also promote the use of research knowledge in health care and produce summaries of evidence-based guidance are "The National Institute for Clinical Excellence" (NICE) and the "Scottish Intercollegiate Guidelines Network" (SIGN).

2.1.4.7 The National Institute for Clinical Excellence (NICE)

The National Institute for Clinical Excellence (NICE) is a part of the NHS established in 1999. It is designed to work with the NHS in England and Wales to appraise health-care intervention. It uses a team of experts who produce guidelines for both the NHS and patients on different medical issues, medical equipment and clinical procedures. The Institute evaluates and rigorously appraises available research findings and formulates and recommends robust and evidence-based guidelines to the NHS (http://www.nice.org.uk).

2.1.4.8 Scottish Intercollegiate Guidelines Networks (SIGN)

SIGN was established in 1993 by the Conference of Royal Colleges and their Faculties in Scotland. Their aim was to develop national guidelines based on a high standard of care which balances the current knowledge and experience against the limitation of available resources. Provision was also made for the placement of mechanisms by which a critical appraisal of broad principles paves the way for implementation of guidelines at a local level appropriate to local needs and constraints.

SIGN consists of multidisciplinary groups; representatives of all the medical Royal Colleges and Faculties of nursing, pharmacy, dentistry, other professions allied to

medicine and patients. SIGN works closely with other national groups and government agencies working in the NHS in Scotland (http://www.sign.co.uk).

The SIGN guideline development methodology involves a rigorous systematic review and appraisal of the existing literature. Its objectives are to sponsor and support the development of a number of evidence-based national guidelines and recommendations which are explicitly linked to supporting evidence. They facilitate their implementation within local practice for the benefit of patients (Harbour and Miller 2001).

Hence NICE and SIGN both develop guidelines in key areas, where sub-optimal clinical care is provided and where there are marked variations in the clinical management which would hamper effective and homogeneous care delivered to patients and adversely influence the outcome of care.

Section 2: Interventions

2.2 Introduction

The assumption that the publication of research findings and evidence in peer review journals would be sufficient and no further action would be needed for this to find its way into clinical practice has been proven to be untrue (Lomas 1993, Effective Health Care 1999). Rapid expanding biomedical knowledge of variable quality, disorganised information resources and information overload indicate the need for more effective strategies for disseminating evidence in clinical practice in order to improve the quality of health care. A wide range of interventions is available, but little is known about the effectiveness of these strategies in daily clinical care (Effective Health Care 1999).

One such strategy is the use of evidence-based guidelines which aim to promote evidence-based practice but in reality the development of valid guidelines does not ensure their use in practice (Feder *et al* 1999). Therefore we need to examine how best to introduce them into clinical practice.

Hence this section of the literature review seeks to provide an overview of the development of evidence-based guidelines and the effectiveness of different strategies for the dissemination and implementation of evidence and a brief account of the barriers to this process.

2.2.1 Evidence-based clinical guidelines

Increasing interest in the use of clinical practice guidelines is stretching across the globe (Haines and Feder 1992). This interest is fuelled by issues that most health care systems face: the large variation in clinical practice with the presumption that at least some of this variation stems from inappropriate care; the intrinsic desire of clinicians to offer, and of patients to receive, the best care possible; and an interest in the management of rising health care costs (Shaneyfelt *et al* 1999, Woolf *et al* 1999, Eccles *et al* 2000).

Clinical guidelines are believed to be a basic tool for establishing optimal clinical practice, for securing the best use of resources and to serve as a valuable means of educational (Mittman 1992, Woolf 1999). The need to employ the evidence-based process to assemble and synthesise the valid and relevant guideline from clinical research has become fundamental for clinical practice during the past decade. Different approaches are used to develop guidelines. They are either produced using evidence-based methodology by a multi-professional group (evidence-based guideline) or by extracting an expert opinion or group consensus (Eccles *et al* 2000).

Although experts may be knowledgeable and experienced, guidelines based on expert opinions are usually unstructured and informal and are open to criticisms of bias and conflicting interests. Guidelines derived from consensus groups are more structured and formal. However research considered may represent a biased sampling and evidence is not generally available for evaluation. Evidence-based guidelines are structured and formal and use rigorous, explicit and reproducible methods to collect and evaluate the

evidence. They are based on systematic reviews and incorporate the values and preferences of patients and clinicians (Jadad 1998).

Clinical guidelines may be developed at different levels, e.g. at a local or national level. However, the locally developed guidelines may be less valid than national ones, as they are more likely to contain bias and are not as well resourced. Nevertheless, the level of development may carry a greater significance, as familiarity with the guideline developers may influence the perceived appropriateness of the guideline and compliance of the targeted group. To overcome this, it is possible that nationally developed guidelines can be further adapted locally and tailored for local use (Feder *et al* 1999).

Guidelines are rapidly outdated, as new evidence becomes available. Therefore the contents are reviewed to incorporate new information and updated on a regular basis by the guideline developers (Haines and Feder 1992, Muir Gray *et al* 1997).

In Scotland the Clinical Resource and Audit Group (CRAG) and the Scottish Intercollegiate Guidelines Network (SIGN) and in England the National Institute for Clinical Excellence (NICE) have been at the forefront of this exercise (see Section 2.1.4.7-8). The SIGN and NICE criteria for appraisal are well known to be founded on high quality research-based evidence (SIGN 50, http:/nice.org.uk/nice-web). As a result, a growing number of guidelines have been developed after exhaustive systematic reviews of the evidence.

2.2.1.1 Effects of clinical guidelines

There are mixed feelings about the potential impact of guidelines (Haines and Feder 1992). Guidelines produced by managers and fund providers may reduce costs and produce good public policies but may be disliked by clinicians and patients as constituting a threat to their personal autonomy or as an unnecessary and/or inappropriate substitute for clinical judgment (Haines and Feder 1992, Mittman *et al* 1992, Woolf *et al* 1999, Haycox *et al* 1999). Furthermore guidelines produced by specialists may seem self-serving, biased and prove impractical in primary care (Haines and Feder 1992, Shekelle *et al* 1999, Woolf *et al* 1999). To specialists, guidelines developed without their contribution fail to meet adequate expertise. Inflexible guidelines with stringent rules about what is appropriate are popular with managers, quality auditors and lawyers but are seen as cookbooks by clinicians who are faced with diverse clinical problems. Nevertheless the main effect of the guideline should be an improvement in the quality of patient care.

Therefore it is important to examine whether clinical guidelines can improve the quality of clinical practice.

This question has been examined by a number of systematic reviews such as those by Grimshaw and Russell (1993), Grimshaw *et al* (1995), the NHS Centre for Reviews and Dissemination (1994) and a study in dentistry by O'Brien *et al* (2000).

In examining the evidence as to whether guidelines can change practice Grimshaw and Russell (1993) conducted a systematic review of 59 well designed, published

evaluations of implementing clinical guidelines to assess the extent of improvements in clinical process or health outcomes. Twenty-four studies investigated guidelines for specific clinical conditions, 27 studied preventive care, and 8 examined guidelines for prescribing or for support services. All but four (55) studies detected significant improvements in the process of care following the introduction of guidelines. Twelve of the 17 studies (Grimshaw *et al* 1995) that assessed the outcome of care reported significant improvements. The authors concluded that guidelines could improve clinical practice when introduced in the context of rigorous evaluations. However, in both studies costs were found not to be carefully summarised.

The NHS Centre for Reviews and Dissemination in 1994 undertook a systematic review to evaluate the effects of guidelines on clinical practice. Eighty-one out of 87 studies examining the effects on the process of care, measured by adherence to recommendations, reported significant improvements and 12 out of 17 studies assessing patient outcomes also reported significant improvement. It was concluded that guidelines are more likely to be effective if they take account of the local circumstances, are disseminated by an active educational intervention and are implemented by patient-specific reminders relating directly to professional activity (Effective Health Care 1994).

A randomised controlled trial carried out by O'Brien and colleagues (2000) evaluated the effectiveness of referral guidelines for the referral of orthodontic patients to consultant orthodontists and specialist practitioners. Several implementation and dissemination strategies were adopted in this study which would be practical for the

general dental services. The participating general dental practitioners received the orthodontic referral guidelines, patients' specific feedback with regard to the outcome of consultation and post-intervention questionnaires by post. Then they were re-issued with the guideline 6 months after the first release accompanied by a survey form about the dentists' perceptions of usefulness of the guideline. The authors concluded that referral guidelines for orthodontic referrals did not influence the behaviour of the general dental practitioners. More research into optimum methods of dissemination and implementation of guidelines for use in general dental services is needed (O'Brien *et al* 2000).

Hence, as stated by Grimshaw and Russell (1993) the successful introduction of clinical guidelines is dependent on many factors but they can improve the quality of clinical care when the clinical context, the methods of development, dissemination and implementation are all appropriate. Nevertheless, whether improvement in the quality of clinical care has been achieved in practice is not very clear, since patients, clinicians, funders and managers define quality differently despite current evidence about the effectiveness and quality of guidelines (Grimshaw and Russell 1993, Grimshaw *et al* 1995 and NHS Centre for Reviews and Dissemination 1994). The evidence on the effectiveness of guidelines is still incomplete and need further investigation (Woolf *et al* 1999).

2.2.1.2 Potential benefits of clinical guidelines

For patients and health care personnel, clinical guidelines can improve health outcomes and the quality of clinical decisions. They offer specific recommendations in doubtful areas of clinical practice, improve the homogeneity of care and provide reliable recommendations that reassure practitioners about the appropriateness of the intervention. Guidelines based on a rigorous appraisal of evidence encourage interventions whose effectiveness has been evaluated and the withdrawal of interventions that are ineffective, dangerous or wasteful by grading the recommendations and documenting the quality of the supporting data.

Guidelines can help influence public policy by calling attention to unrecognised health problems, possible preventive interventions, neglected patient populations and high-risk groups. Services that were not previously offered to patients may be made available as a response to published guidelines.

Clinical guidelines can support quality improvement activities such as audits and reminder systems as they constitute a common point of reference and can be useful for medicolegal protection or for setting practice policies.

Finally, guidelines can be effective in improving the efficiency of health care by standardising care, reducing prescriptions and interventions. Consequently, they can optimise value for money by releasing the health system resources needed for other health care services.

2.2.1.3 Potential problems with clinical guidelines

The most important limitation of guidelines is that recommendations can be wrong or at least can be so for an individual patient. This can occur through unintentional oversights by busy or weary members of a guideline developer group, lack of scientific evidence, the poor composition of the guideline developer group or by recommendations influenced by opinions and anecdotal clinical experience.

A flawed clinical guideline has potentially a direct impact on patient care. Unsound clinical guidelines provide inaccurate recommendations and can encourage sub-optimal, ineffective or harmful practices.

When guidelines are evidence-based, they can be found to be difficult to use, impractical, time-consuming or inconvenient (Cabana *et al* 1999).

Generalised recommendations as opposed to a menu of options or recommendations for mutual decision-making may overlook patients' preferences and be inappropriate for an individual's care (Woolf 1997). Guidelines that are not flexible cannot be customised to an individual patient's predicaments (Cabana *et al* 1999).

The main benefit of clinical guidelines, resulting in providing a coherent practice and reduced disparity of care, may come at the cost of reducing customised care for patients with special needs. Although, according to the definition of a clinical guideline this should not be the case.

Differing guidelines from different professional bodies can also be confusing, frustrating and biased (Feder 1994, Haynes and Haines 1998, Cabana *et al* 1999). Outdated recommendations may encourage obsolete practices and technologies. Perhaps it would be necessary for recommendations to have a "use-by" date or be updated regularly to reflect new evidence (Haines and Feder 1992, Shekelle *et al* 1999, Muir Gray 2001).

Clinical guidelines can also damage clinicians professionally, as auditors may unfairly criticise the quality of care provided, based on a recommendation from unsound guidelines or by not understanding why sound guidelines have not been followed.

A negative or neutral recommendation may lead providers to discontinue a particular service which may be useful for individual patients. If guidelines state that there is no evidence for the effectiveness of an intervention, this may prompt the funding bodies to withdraw funds or stop investing in further research in these areas (Haycox *et al* 1999). Recommendations for costly interventions may shift limited resources from services of great importance to some patients (Haycox *et al* 1999). In addition, the tendency of guidelines to focus on specific health issues can give misleading impressions to the public and to health providers about the relative importance of a disease or the effectiveness of interventions.

Some clinical guidelines, especially those developed by groups unconcerned with financing, may encourage expensive interventions that are unaffordable or use resources which are meant to be for more effective services (Birch *et al* 1995, Haycox *et al* 1999).

Therefore clinical guidelines must be only one of the options for improving the quality of care and should be developed from a patients' perspective while at the same time considering the needs of the whole community (Haycox *et al* 1999, Woolf *et al* 1999). They are not laws and hence should not be followed rigidly like a recipe book but should weigh the risk and benefits of evidence with clinical experience and the individuality of clinical circumstances (Bader *et al* 1999, Hurwitz 1999, Batchelor 2000).

2.2.1.4 Developing guidelines

Guidelines can be developed for a wide range of subjects. Potential issues can arise from "the evaluation of the major causes of morbidity and mortality for a given population" or "the uncertainty about the appropriateness of healthcare practice" or "the effectiveness in improving patient outcomes" or "the need to conserve resources in providing care" (Muir Gray *et al* 1997, Shekelle *et al* 1999). Often the selected subject requires refinement before the evidence can be appraised. This is usually carried out by discussion amongst clinicians, patients and the potential users or evaluators of guidelines. Then it is necessary to set up multidisciplinary guideline developers' groups which consist of "a management team" responsible for the day to day running of the work, such as the identification, synthesis and interpretation of relevant evidence and a "guideline development group" in charge of producing recommendations on the basis of existing evidence or its absence (Grimshaw and Russell 1993, Shekelle *et al* 1999).

When the sources of evidence are identified, they are assessed by performing a systematic literature review. The purpose of a systematic review is to collect all the

available evidence, evaluate the methodological quality, appraise the potential relevance to the clinical question under deliberation and summarise the findings (Mulrow 1994, Shekelle *et al* 1999, Harbour *et al* 2001).

The evidence, once collected and summarised, is translated into clinical recommendations and health policies. Opinions are used to interpret evidence and to derive recommendations in the absence of direct evidence. This is needed to assess issues such as the generalisibility of evidence, for example, to assess to what extent evidence from small randomised clinical trials or controlled observational studies can be generalised (Muir Gray *et al* 1997, Shekelle *et al* 1999). Besides, the strength of the evidence offered by an individual study depends on the ability of its design to minimise the possibility of bias and maximise attribution of the evidence on which the recommendation is based. For example, recommendations based on clinical judgement and experience are likely to be more susceptible to bias and self-interest (Harbour *et al* 2001).

Accordingly, evidence is categorised to reflect its susceptibility to bias or confounding factors that reduce its reliability. This is a shorthand method of conveying specific aspects of the evidence to any reader of the guideline. Therefore it is important to grade each recommendation in the guideline to indicate the guideline development group's confidence that use of the guideline will produce the desired health outcome.

A simple scheme for the classification of the evidence that supports statements in guidelines and the strength of the recommendations, is based on the work of the US

Agency for Health Care Policy and Research (AHCPR) (US Agency for Health Care Policy and Research 1993, Muir Gray *et al* 1997, Shekelle *et al* 1999, Harbour *et al* 2001).

The simple scheme for the classification of evidence that supports statements in guidelines and the strength of recommendation is set out as follows (Shekelle *et al* 1999):

2.2.1.5 Classification schemes

Category of evidence:

Ia- evidence from meta-analysis of randomised-controlled trials

Ib- evidence from at least one randomised-controlled trial

IIa- evidence from at least one controlled study without randomisation

IIb- evidence from at least one other type of quasi-experimental study

III- evidence from non-experimental descriptive studies, such as comparative studies, correlation studies, and case-control studies

IV- evidence from expert committee reports or opinions or clinical experience of respected authorities, or both

Strength of recommendations:

- Directly based on category I evidence
- Directly based on category II evidence or extrapolated recommendations from category I evidence

- Directly based on category III evidence or extrapolated recommendations from category I or II evidence
- Directly based on category IV evidence or extrapolated recommendations from category I, II or III evidence.

This was the system proposed by AHCPR and initially used by SIGN. However, the weaknesses of the existing system became apparent following several years of guideline development by SIGN. In 1998 they began to review and, where appropriate, to refine the system for evaluating guideline evidence and the grading recommendations. For example, the previous grading system was designed for the application of questions of effectiveness where randomised controlled trials are accepted as the most robust study design with the least risk of bias in the outcome. In many areas of clinical practice RCTs may not be practical or ethical to undertake and for many questions other types of study design may provide the best evidence.

Often too, guideline users are not clear about the implications of the previous grading system. They misinterpret the grade of recommendation as relating to its importance, rather than to the strength of the supporting evidence and could then possibly fail to give due weight to a low-grade recommendation (Harbour *et al* 2001).

Therefore the revised grading system by SIGN is intended to strike a balance between incorporating the complexity of the type and quality of the evidence and maintains clarity for guideline users (Harbour *et al* 2001). The differences from the AHCPR system are that the study type and the quality rating are combined in the evidence level;

the grading of recommendations extrapolated from the available evidence is clarified; and the grading of recommendation is extended from three to four categories.

The revised grading system for recommendations in evidence-based guideline by SIGN is shown below (Harbour *et al* 2001).

Level of evidence

1++ High quality meta-analyses, systematic reviews of RCTs, or RCTs with a very low risk of bias.

1+ Well-conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias.

1- Meta-analyses, systematic reviews of RCTs, or RCTs with a high risk of bias

2++ High quality systematic reviews of case-control or cohort studies or high quality case-control or cohort studies with a very low risk of confounding, bias, or chance and a high probability that the relationship is causal

2+ Well-conducted case-control or cohort studies with a low risk of confounding, bias, or chance and a moderate probability that the relationship is causal

2- Case-control control or cohort studies with a high risk of confounding, bias, or chance and a significant risk that the relationship is causal

3 Non-analytic studies, e.g. case reports, case series

4 Expert opinion

Grades of recommendations

A At least one meta-analysis, systematic review, or RCT rated as 1++ and directly applicable to the target population *or*

A systematic review of RCTs or body of evidence consisting principally of studies rated as 1+ directly applicable to the target population and demonstrating overall consistency of results

B A body of evidence including studies rated 2++ directly applicable to the target population and demonstrating overall consistency of results *or*

Extrapolated evidence including studies rated as 1++ or 1+

C A body of evidence including studies rated 2+ directly applicable to the target population and demonstrating overall consistency of results *or*

Extrapolated evidence including studies rated as 2++

D Evidence level 3 or 4 *or*

Extrapolated evidence from studies rated as 2+

In the end, the guideline usually receives an external review to ensure content validity, clarity and applicability (Shekelle *et al* 1999). Since guidelines cannot be static and the systematic review supporting the guideline is continually updated to reflect the new available evidence (Haines and Feder 1992, Shekelle *et al* 1999, Muir Gray 2001).

2.2.1.6 Guidelines in dentistry

Incorporation of evidence-based dental research findings into the care of patients has been found to be central in maximising the benefit and reducing any harm of dental treatment (Clarkson *et al* 1999). However, the development of evidence-based guidelines in dentistry is in the early stages. A review in 1995 of guideline development by various dental organisations and specialists in the United States revealed a lack of systematic reviews and a reliance on expert opinion acquired through unstructured and

untested methods of consensus. They found that there are relatively few research findings based on randomised controlled trials and other treatment outcome studies in dentistry that have evaluated clinically relevant interventions (Bader and Shugars 1995). Where evidence is available, the range of outcomes used to evaluate treatment is usually limited to technical issues such as mechanical features and survival of teeth and restorations (Bader *et al* 1999).

Hence, in the last decade, steps have been taken to improve the process of preparing, gathering and implementing the best research findings in clinical dental practice (Bader *et al* 1999, Sutherland 2000). As a result a number of guidelines have been developed in the UK in different specialities in dentistry to assist with the achievement of an optimal clinical practice. The number of guidelines in different aspects of dentistry are slowly growing and a few examples of published national guidelines in recent years are Dental Radiology (1998), SIGN Guideline in targeted caries prevention in 6-16 year olds attending for dental care (2000), Clinical examination and record keeping good practice guideline (2001).

Similarly a number of guidelines on the topic of the management of unerupted and impacted third molar teeth have been developed, including the document produced by a working party of the Faculty of Dental Surgery of the Royal College of Surgeons of England (1997) and the NICE Guideline (1999). The most recent guideline in this subject was the SIGN Guideline for the management of unerupted and impacted third molar teeth (2000).

The SIGN guideline on " Management of Unerupted an Impacted Third Molar Teeth" is intended to serve as a tool by which scientifically valid and reliable standards of clinical management of unerupted and impacted third molar teeth can be implemented (SIGN 43, 2000). The SIGN Guideline may well result in a reduction in the variation of current clinical management of such teeth and facilitate appropriate and effective clinical care, in addition to containing costs. As suggested by Shepherd (1993) conservative treatment with more rigorous adherence to specific recommendations for removal of third molar teeth could reduce surgical cases by up to 60%.

The present SIGN guideline (43, 2000) provides valid recommendations about the risks and benefits of treatment and incorporates management strategies for the removal of third molar teeth which can then be used to facilitate clinical decision-making.

2.2.2 Interventions for introducing evidence-based guidelines in daily clinical practice

There are three ways of incorporating evidence, including evidence-based guidelines, into clinical practice. These are "diffusion", "dissemination" and "implementation".

Lomas (1993) describes diffusion as a passive concept that is unplanned and uncontrolled, in which untargeted information flows away from its origin. This process depends on its audience's interest, motivation and their effort to seek the information.

Dissemination, which literally means ''to scatter, or sow'' (Concise Oxford Dictionary 1997), is a rather more active process. It is defined as the launching of targeted and tailored information aimed specifically at a particular audience which raises awareness of new and relevant research knowledge. Dissemination alone is not sufficient enough though to promote changes in practice (Lomas 1993). This has been shown in a review of 19 studies of passive dissemination/diffusion of consensus-derived recommendations for practice. They concluded that there was little evidence that passive dissemination/ diffusion alone resulted in behaviour change (Lomas 1993, I).

Implementation involves identifying and assisting in overcoming the barriers to utilise the knowledge obtained from a tailored message, prompting the clinician to take on the evidence. "It is a more active process still which uses not only the message itself but also organisational and behavioural tools that are sensitive to limitations and abilities of identified clinicians in identified settings" (Lomas 1993).

Implementation is usually distinguished from dissemination strategies in that implementation aims to achieve change in the recipient's action and behaviour.

In practice, dissemination and implementation activities are closely related and they exist on a continuum because sufficient awareness and understanding is usually an essential prerequisite for changing behaviour and practice (Marriott and Palmer 2000). Strategies for dissemination and implementation must be maintainable and dynamic, taking into account changing evidence and their own effectiveness (Haines and Jones 1994).

2.2.2.1 Effectiveness of different interventions strategies

Considerable interest has been generated in developing and implementing research findings and clinical practice guidelines. Nevertheless, little is known about the effectiveness of implementation of evidence or in other words, "What makes an intervention successful in routine health care?" However, a number of well-planned systematic reviews have attempted to evaluate the effectiveness of different types of interventions and they have usually used any changes in the target group's performance and behaviour as an outcome measurement. Alternatively, where strategies support awareness of a clinically effective practice, they have measured improvements in health outcomes (Marriott and Palmer 2000).

There are different types of intervention for incorporating evidence into practice and they can be categorised into broad strategies such as "continuing medical education" or specific interventions such as educational materials, outreach visits, local opinion

leaders, patient-mediated interventions, audit and feedback, reminders, peer review and multifaceted interventions.

2.2.2.1a Continuing medical education

Haynes *et al* (1984) carried out a systematic review of 248 studies (including 35 randomised-controlled trials) evaluating the effectiveness of continuing medical education. However, only 7 studies met their inclusion criteria, of which 6 observed improvement in the performance of clinicians. One out of 3 studies measuring patient outcome observed statistically significant improvements. Interventions commonly evaluated in this study were "educational materials", "audit and feedback" (A&F) and "opinion leaders".

The authors concluded that the studies were encouraging and practitioner behaviour could be improved by continuing medical education interventions (Haynes *et al* 1984).

Bero and his colleagues (1998) investigated a series of systematic reviews of interventions that presented a change of professional performance or treatment outcome, but only 18 reviews met their inclusion criteria. They were categorised as continuing medical education, audit and feedback, computerised decision-support systems, or multifaceted interventions. Most of the reviews identified modest improvements in performance after the interventions. However, passive dissemination of information was generally ineffective in altering practice, no matter how important the issue or how valid the assessment methods were (Bero *et al* 1998).

Davis *et al* (1999) undertook a systematic review of 64 studies. Only 14 studies, which included 17 interventions, met their inclusion criteria. Studies included were randomised-controlled trials of formal didactic and/or interactive continuing medical interventions (conferences, courses, meetings, symposia, lectures and other formats). Nine of the 17 interventions generated positive changes in professional practice and 3 out of 4 interventions altered health care outcomes in one or more ways. The authors concluded that interactive continuing medical education provides the practical skills that can effect change in professional practice and health care outcomes and that didactic sessions by themselves do not appear to be as effective in changing physician performance or improving patient care. However, the limited number of randomised-controlled trials and settings may mean that these findings do not lend themselves to generalisation. However, the fact that this small number of studies identified significant results may imply large and fairly consistent effects from interactive educational methods (Davis *et al* 1999).

2.2.2.1b Mass media and printed educational materials

Grilli and his colleagues (2002) investigated the impact of mass media (i.e. radio, television, newspapers, magazine, leaflets, posters and pamphlets) on utilisation of health services by undertaking systematic reviews of 69 studies which met their inclusion criteria. Although the reviewers were confronted with methodological flaws in the design of relevant trials, they found that all the studies apart from one showed that mass media were effective. The authors concluded that those engaged in promoting a better uptake of research information in clinical practice should consider mass media as a tool that may encourage the use of effective services and discourage those of

unproven effectiveness.

A simple provision of information was examined in a systematic review of 11 studies concerned with dissemination of printed educational materials. These were compared with dissemination with no active intervention, as well as with enhanced intervention combined with an educational or marketing strategy such as audio-visual materials or electronic publications. The effects were small but increased when additional strategies were added. Of these, individual educational outreach sessions and the use of local opinion leaders as agents of dissemination appeared promising (Freemantle *et al* 1999).

Soumerai *et al* (1989) reviewed interventions to improve drug prescribing in primary care. Forty-four studies met their inclusion criteria and 85% of inadequately controlled studies reported positive findings, compared to positive results in 55% of well-conducted studies. They concluded that the dissemination of educational materials including guidelines might change knowledge or attitudes; nevertheless it has little or no noticeable effect on the actual prescribing. However, they noted the distribution of educational materials was relatively inexpensive and may be worthwhile even if it only results in small changes in practice.

2.2.2.1c Outreach visit

Outreach (academic detailing) visiting is a complex intervention using trained personnel or an educator who meets with clinicians in their practice settings to provide information. The information given may include feedback on the clinicians' performance. One of the components of this method is the participation of clinicians in

conferences, lectures, workshops, or training sessions outside their own practice settings. Participation in conferences can be in the context of a small group in an active capacity, or in a big group in a passive capacity.

Newton-Syms *et al* (1992) studied the effectiveness of a single outreach visit and educational materials, compared to a no-intervention control group of general practitioners in the United Kingdom, in an effort to encourage rational prescribing of non-steroidal anti-inflammatory agents. They reported that the intervention was effective. Overall, there was a decrease in the percentage of prescribing costs in the intervention group.

Thomson O'Brien *et al* (2002, III) conducted a review of 18 studies evaluating the effectiveness of educational outreach visits. In 12 of the 13 trials of combined interventions, there were positive effects in favour of the intervention group (15-68% relative improvement). Three trials reported significant improvement where outreach visits alone were compared with a no-intervention control group (24%-50% relative improvement). One trial found outreach visits to be more effective than audits and feedbacks. Another observed outreach visits using patient-related content (i.e. case study) to be more effective than using performance summaries for content (i.e. statistical information). Only one trial reported that the effect of outreach visits decreases over time. The authors concluded that this type of intervention is a promising approach for modifying professional behaviour, particularly when combined with additional interventions, as its effects are small to moderate. However the cost-effectiveness of this approach in different circumstances and health settings is unclear

and there is also the need to monitor the long-term performance of the effectiveness of outreach visits (Thomson O'Brien *et al* 2002,no III). We should also consider that most of this research is from North America and its general application to the UK setting is not certain.

A review of 50 randomised-controlled trials of continuing education suggested that strategies which incorporate feedback on performance, the involvement of learners in setting priorities, or face-to-face encounters between clinicians and an educator could be very effective (Davis *et al* 1992). However, only a minority of these studies have assessed the impact on patient outcomes. Strategies linked to activities that facilitate or reinforce practices consistently improve the performance of clinicians.

2.2.2.1d Local opinion leaders

Once an opinion leader who is 'educationally influential' in a given system adopts an innovation, his colleagues may take this on board. Thomson O'Brien *et al* (2002, IV) systematically reviewed eight studies involving more than 296 health professionals. They found that using local opinion leaders resulted in mixed effects on practice and that further research is needed before widespread use of this intervention could be justified.

2.2.2.1e Patient-mediated interventions

Patient-mediated interventions are where the provision of information or support to patients is used to indirectly influence the performance of clinicians (Haines and Jones 1994).

Educating patients about the effectiveness of interventions is an attempt to change the behaviour of professionals. Giving patients more information about the probabilities of different outcome of the treatment and their impact on their quality of life may influence the treatment that patients choose. Information is provided for patients by direct mailings, patient counselling delivered by others, materials given to patients or placed in waiting rooms or the use of interactive computer software programmes.

In a systematic review carried out by Davis and colleagues in 1995, patient-mediated interventions appeared to improve the provision of preventive care (Davis *et al* 1995).

2.2.2.1f Audit and feedback

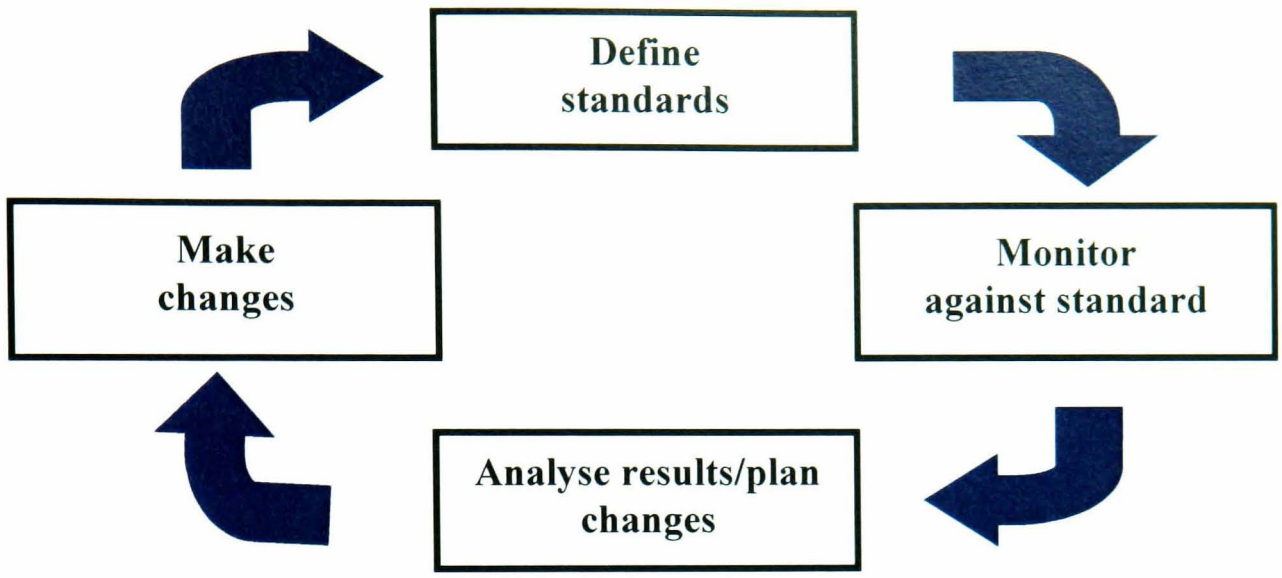
The 1989 government White Paper (Department of Health Working Paper 6,1989), ''Working for Patients'', introduced the concept of medical audit as a method of improving the quality of patient care and continuing the medical education aimed at improving a health care providers performance (Johnston *et al* 2000).

Audit is "the systematic critical analysis of the quality of medical care, including the procedures used for diagnosis and treatment, the use of resources and the resulting outcome and quality of life for the patient" (Crombie *et al* 1993). Information feedback is defined as the use of comparative information from statistical systems (Thomson O'Brien *et al* 2002, no II) and the supply of positive and negative information on (gaps in) performance (Grol 1992).

An audit provides information and a summary of the clinical performance in health care over a specified period of time and may include recommendations for clinical action. Information on clinician performance may have been obtained from records, computerised databases, observation, or from patients (Thomson O'Brien *et al* 2002. I). An audit is a dynamic, cyclical process (Figure 2.1: an audit loop) and its starting point is the assessment of the quality of current health care. Deficiencies in health care can then be identified by comparing the current practice with a set standard and implementing the changes to improve the delivery of health care and the quality of care reassessed. This closes the audit loop and the procedure begins again. As the first attempt at change is often only partially successful in resolving the deficiencies and to achieve the desired level of care, it may be necessary to re-enter the cycle and repeat the steps (Crombie *et al* 1993, Ch: 2).

The use of a defined set of standards is essential for any audit and the value of a set standard depends on the way it is framed. Audit can be classified into an ''internal audit'', i.e. audit performed by the providers themselves and an ''external audit'', i.e. the providers getting data on their performance from others. The evolution of standards over successive audits towards improved levels of care is essential and the standard should be agreed by a comparison with the existing evidence, or by a critical assessment of current practice. It should be set by discussion among all involved so that the standard is one which everyone wants to achieve, rather than one imposed by others.

Figure 2.1 The audit loop



A number of studies evaluate the effectiveness of audit and feedback. One such trial was a cluster randomised controlled trial by Eccles and colleagues examining the effect of audit and feedback on primary care radiology referrals. Their findings suggested that A&F does not change behaviour for primary-care radiology referrals (Eccles *et al* 2001).

Also to be considered is a review of 13 trials assessing the effects of audit and feedback on the practice of health professionals and patient outcomes by Thomson O'Brien and colleagues (2002, I). They compared an audit and feedback group to a no-intervention control group and 8 studies reported statistically significant changes in favour of the experimental group on at least one major outcome but the effects were generally only small to moderate (-16% to 152%). However, they had serious concerns about the studies that they examined. Many of their concerns were about inappropriate levels of randomisation or analysis which would cause an overestimation of effect size. The review concluded that A&F could be effective in improving performance, in particular

for prescribing and diagnostic test ordering. The Audit has also been shown to be effective in monitoring practice performance, resulting in changing medical practice (Grol 1992). However, Thomson O'Brien and colleagues (2002, I) do not support the widespread use of A&F, since their findings suggested a great uncertainty about the efficacy of this approach.

A systematic review of 4 trials by the same authors that were originally carried out in 1997 (and updated regularly since by Cochrane reviewers) (2002, II) compared the effects of A&F with other interventions in changing health professional practice. They found little evidence of the measurable effect of adding a complementary intervention such as a local consensus process to audit and feedback compared to audit and feedback alone.

2.2.2.1g Reminders

A "reminder" is defined as any intervention, manual or computerised, that prompts the clinician to remind them of desired actions for individual patients. They prompt the clinicians to recall information that they already know or would be expected to know by presenting information in a different, more accessible or relevant format at an appropriate time.

An example of a "reminder" is a computerised decision support system that incorporates information from research. This system uses an active knowledge database which generates case-specific advice at the time of consultation. Inter-visit reminders

are targeted at clinicians between visits when there is evidence of sub-optimal care for specific patients.

Examples are enhanced laboratory reports which are lab reports following any abnormal result, targeted at clinicians, which includes additional information about specific follow-up recommendations or implicit reminders which include predictive values for abnormal test results without an explicit recommendation for action.

A systematic review of relevant trials focused on computer-generated reminders showed that such systems can improve the performance of practitioners in terms of decisions on drug dosage, the provision of preventive care and clinical management of patients but not in the diagnosis. More research is also needed on patient outcomes (Johnston *et al* 1994).

Similarly, use of computer-based decision support systems has led to improvements in decisions on drug dosage (9 of 15 studies), provision of preventive care (14 of 19 studies), practitioner performance (19 of 26 studies) and patient outcomes (4 of 7 studies) but not in diagnosis, as shown by Hunt *et al* (1998) in a review of 68 studies.

Another review of 98 studies by Balas *et al* (1996, I) on computerised information systems found that different interventions, including provider and patient prompts, computer-assisted patient education and computer-assisted treatment planners improved patient care.

Goodey and colleagues conducted a randomised controlled trial examining three primary to secondary referral strategies for patients with third molar problems (current practice, clinical algorithm and the neural network-based computer programme). They concluded that the clinical algorithm (flowchart) was the most appropriate option for a reminder strategy. The neural network-based computer programme (the third molar decision support system) performed less well than either current practice or the clinical algorithm (Goodey *et al* 2000).

A published study in dentistry by Kay and colleagues on an evaluation of CAL in developing clinical decision-making skills showed that educational interventions, such as CAL, had no effect on dentists' treatment decision-making behaviour (Kay *et al* 2001).

Further, in a systematic review of 4 trials to assess the effects of audit and feedback, there was a comparison with other interventions such as reminders in changing health professional practice by Thomson O'Brien and colleagues (2002, II). Two of three trials that compared audit and feedback with reminders reported that reminders were more effective in improving the delivery of some preventive services but the results were not striking.

Similarly a randomised controlled trial by Eccles *et al* (2001) investigating the effect of reminder messages on primary care radiology referrals showed that the routine attachment of educational reminder messages (manual paper reminder) to radiographs is

effective and can reduce the number of requested radiographs by 20% (Eccles *et al* 2001).

2.2.2.1h Peer review

From a review of 12 evaluations of physician profiling, defined as peer comparison feedback, 10 studies observed statistically significant improvements on various clinical procedures but the effects were small (Balas *et al* 1996, II). The review concluded that peer comparison alone is unlikely to result in substantial quality improvement or cost control, and may be inefficient.

2.2.2.1i Multifaceted interventions

This is a combined strategy which can include a number of interventions and various reviews support the superior efficiency of multifaceted interventions compared to single intervention.

A review of 61 studies on the effectiveness of implementing guidelines in primary care, by Wensing *et al* (1998) showed that strategies combining more interventions tend to be more effective but may incur greater costs.

Davis *et al* (1995) undertook a systematic review of 99 randomised trials of continuing medical education strategies. They evaluated 160 individual comparisons, as some had more than one comparison - such as three arms randomised trials comparing two active interventions with control. Ninety-nine out of 160 (62%) showed improvement in at least one major outcome in clinician performance or patient outcome. Thirty-three

(33%) showed no improvement and 8 (5%) showed mixed results. The review compared the effectiveness of single or multifaceted interventions: 81 single intervention strategies were used in the trials, 49 (60%) of these demonstrated improvements; 39 interventions used two educational methods of which 25 (64%) were positive; 39 interventions used three or more educational strategies, with 31 (79%) of these being positive.

The following single intervention strategies were generally effective: educational outreach, opinion leaders, patient-mediated interventions and clinician reminders. More variable results were found for the following strategies: A&F and short formal continuing medical education programmes (such as conferences). Multifaceted interventions were found to be more successful. Authors concluded that widely used continuing medical education methods such as conferences have little direct impact on improving professional practice. More effective methods such as systematic practice-based interventions and outreach visits are seldom used by continuing medical education providers.

Oxman *et al* (1995) in their review identified studies involving 12 comparisons of educational materials, 17 of conferences, 4 of outreach visits, 6 of local opinion leaders, 10 of patient mediated interventions, 33 of audit and feedback, 53 of reminders, two of marketing, 8 of local consensus process and 15 of multifaceted interventions. All the interventions showed some effect at least some of the time. However, even relatively complex and intensive interventions, such as outreach visits and the use of opinion leaders, have at best a limited effect.

Hulscher and colleagues (2002) carried out a systematic review of 55 studies with 83 comparisons between intervention and control groups. They showed that most interventions were found to be effective in some studies but not in others. Fourteen comparisons of multifaceted interventions versus no intervention showed an absolute change of preventive services varying between -3% and +64%. Six comparisons of multifaceted interventions versus group education reported absolute changes varying between -31% and +28%. All these comparisons used randomised groups. Reviewers concluded that there is no solid basis for assuming that a particular intervention or package of interventions will work. Effective interventions to increase preventive activities in primary care exist but there is a considerable variation in the level of change achieved with effect sizes usually small or moderate. Multifaceted interventions may be more effective than single interventions because more barriers to change can be addressed (Hulscher *et al* 2002).

In conclusion a summary of the relative effects of these various strategies is assessed by Bero and colleagues (1998) and is shown in Table 2.1.

Therefore it seems a combination of various strategies is more successful in altering practice. However, even multifaceted interventions do not always improve performance (Bero *et al* 1998). There are no magical ways for improving the quality of health care but a wide range of interventions is available which, if used appropriately, can lead to substantial improvements in the application of research and, ultimately, the effectiveness of health care (Oxman 1995).

Table 2.1 **Effects of interventions to promote behavioural change among health professionals**

Consistently effective interventions

- Educational outreach visit
- Reminders
- Multifaceted interventions
- Interactive educational meetings (participation of health providers in workshops encouraging discussion and practice)

Interventions of variable effectiveness

- Audit and feedback
- The use of local opinion leaders
- Local consensus process

Interventions that have little or no effect

- Educational materials
- Didactic educational meetings
- Patient mediated interventions (feedback from patients incorporated into intervention)

Bero *et al* 1998, Effective Health Care 1999

2.2.2.2 Implementation of evidence

When it comes to the implementation of a change as has been previously discussed, "none of the approaches is superior and there are no magic bullets'' (Oxman 1995). Different research findings may require different methods of implementation and different groups of individuals may experience different obstacles to change. Equally, implementation of clinical guidelines and introduction of change are complex process and do not usually involve a single action but a well-planned and organised stepwise process, linked to the specific barriers to change (Grol 1997, Hulscher *et al* 2002). To promote the uptake of evidence and adaptation of guidelines, these obstacles and also the strategies to overcome them should be identified (Haines and Donald 1998). To facilitate this process a five step model, around which implementation of evidence in clinical practice can be designed, was suggested by Grol (Grol 1997):

1) A definite plan for changing clinical practice should be developed in a way that this guideline or change proposal is based on clinical expertise and scientific evidence.

2) The obstacles and facilitators to a change should be identified before an intervention or combination of interventions is selected for implementing a change.

3) Interventions should be linked to the obstacles. Since clinicians may experience a variety of barriers, often different interventions are needed which are linked to different phases in the process of change.

4) A plan should be devised in such a way that allows concrete tasks to be divided and a time schedule to be set up. Usually it is neither desirable nor feasible to use all interventions at once, so a division of small-scale activities and evaluations is a more appropriate action.

5) The progress should be evaluated when different steps of the plan are carried out. The plan should be adapted if the results of the evaluations show this is necessary.

Implementation strategies may take account of such models but it seems there is not enough knowledge to support or reject this model. However, it is clear that research in implementation of guidelines in daily clinical practice should be complemented by an evaluation of barriers which hinders the implementation of a change. This process then may facilitate targeting interventions to specific barriers (Mittman 1992, Freemantle 1995, Smith 1995, Grol 1996,1997).

2.2.2.3 Barriers and Facilitators and Behavioural Theories

Barriers and facilitators are defined as factors that prevent or enhance changes in clinical behaviour. The barriers to the dissemination and implementation of research findings in making decisions about health care are multifaceted. Despite attempts to improve the quality of evidence, adaptation and utilisation by clinicians has been less than optimal and it seems that the successful implementation of research findings into practice is determined by many factors (Table 2.2). Research focused upon the medical profession has shown that the barriers and facilitators may be related to the clinicians'

knowledge, skills and attitudes or to the social and community environments, organisation, economic and legal context in which clinicians work (Mittman *et al* 1992, Lomas 1993 II, Haines and Donald 1998, Cabana *et al* 1999).

Potential barriers to physicians' guideline compliance were investigated in a review of 76 articles including 120 different surveys (Cabana *et al* 1999). In this review, 84% of the participants identified "lack of awareness", 89% "lack of familiarity", 91% "lack of agreement", 65% "lack of self-efficacy", 90% "lack of outcome expectancy" (*i.e.* the expectation that a given behaviour will lead to particular consequences) and 65% "inertia of previous practice" as potential barriers to the adherence to the guidelines. Similarly 10% of the participants indicated the "guideline related factors" (*i.e.* not easy to use, cumbersome, confusing) and 10% "patient related factors" (*i.e.* the inability to reconcile patient preference with guideline recommendations) with 10% "environmental related factor" as a possible barrier (Cabana *et al* 1999). This study indicated that clinicians' lack of awareness of or unfamiliarity with the existence of new evidence or guidelines, knowledge about their own practice being different from the recommendations, or attitudes of clinicians towards evidence, as well as attitude towards the producer, promoter and content of the guideline in terms of their usefulness and reliability, may all influence the outcome of an implementation strategy. Compliance with the guideline recommendations requires sufficient individual skills, team expertise and organisational competence. Inadequate skills of individual clinicians in searching and locating the information about the evidence could result in sub-optimal compliance and practice.

Social factors too can act as a barrier or facilitator for clinical practice. For example, colleagues at the practice, managers, opinion leaders, patients or other members of a professional team or the authorities can be supportive of or resistant to recommendations for a change (Haines and Donald 1998).

Organisational factors, which include practicality within an existing practice, setting or routines; availability of guidelines in the workplace; and local infrastructure or rules, are also considered as potentially significant obstacles or facilitators (Haines and Donald 1998). Economic factors such as the availability of or lack of time and personnel resources, secondary costs in implementing the change and financial incentives or sanctions have also been cited as potential barriers or facilitators.

Table 2.2 Potential barriers to changing clinical practices

Knowledge and attitude

- Information overload
- Clinical uncertainty
- Influence of opinion leaders
- Obsolete knowledge

Patient factors

- Demands for care
- Perceptions and beliefs about appropriate care
- Compliance with clinical guidance

Practice environment

- Time constraints
- Poor practice organisation

Educational environment

- Outdated undergraduate education
- Inappropriate continuing education
- Lack of incentives to participate in effective educational activities

Wider health system

- Inappropriate funding system
- Lack of financial support for innovation
- Failure to provide practitioners with access to appropriate information

Social environment

- Media influence in creating demands for treatment
- Commercial concerns promoting products and equipment

Oxman and Flottrop 1998, Haines and Donald 1998
(Adopted from McGlone *et al* 2001)

Unfortunately, insight into barriers and facilitators for change is often incomplete. Even if knowledge of these elements is available, their value for achieving change in a specific clinical setting is limited because these factors may be specific for those clinicians or that setting (Grol 1997, Cabana *et al* 1999).

In dentistry, very few studies have investigated the factors influencing change in dental practice and the likely barriers to change towards an evidence-based approach (Hunt *et al* 1983, Kay and Blinkhorn 1996, Chestnutt and Kinane 1997, Rushton *et al* 1996, McColl *et al* 1999). Nevertheless the factors influencing dental practice reflect the potential barriers to change similar to those assessed in the medical literature (McGlone *et al* 2001). Factors such as patients' opinion and values and dentists' knowledge, attitude, confidence and conscience are reported to be important factors governing the philosophies of treatment (Kay and Blinkhorn 1996). For example, patients can influence specific treatment decisions such as periodontal (Chestnutt and Kinane 1997) and endodontic treatments (McColl *et al* 1999). Patient opinion and medicolegal factors can effect the dentists' decision about practices of taking bitewing radiographs (Rushton *et al* 1996). Attitude and knowledge of sealants and preventive dentistry have also effected preventive treatment decisions (Hunt *et al* 1983, Main *et al* 1997). Organisational and funding arrangements, such as the NHS fee scale, have been found to be a major factor in the adaptation of new techniques (McColl *et al* 1999).

To have a better understanding of these factors it is essential to study the underlying behavioural theories. A range of barriers and facilitators for change is suggested by different theories on behavioural change (Grol 1997, Haines and Donald 1998, Effective

Health Care 1999, Feder *et al* 1999, Marriott and Plamer 2000). Three theories on behavioural change are very briefly explained here.

Cognitive theories include those focusing on rational information seeking, decision-making and management theories that highlight organisational conditions necessary to improve care. These models see factors such as belief, attitudes and intentions as central influences in changing clinicians' behaviour. A refinement of these models is the stage model of behaviour which explains the factors thought to influence change in different settings. Individuals are thought to go through a series of stages and for each stage different interventions are needed. One stage model suggests that the different stages are "pre-contemplation", "contemplation", "preparation", "action" and "maintained" behaviour. It is important to target a specific group and discover its needs, barriers and drives for change at each stage. Another model is based on receptivity, to sort classified individuals into "innovators," "early adaptors," "early majority," "late majority" and "laggards." As groups differ in the degree to which they are willing to change or perceive different benefits and barriers, different interventions and techniques are essential.

Learning theories offer an explanation of how behaviour is determined by modifying factors which control behaviour, such as rewards (positive consequences) and punishments (negative consequences) which are imposed from the outside on to an individual. The effectiveness of these reinforcing factors depends on how desirable they are and how motivated the individuals are to obtain them. Interventions such as the reminder system, audit and feedback and fee-for-service are based on this theory. It is

unlikely that such approaches would work in all circumstances. However, reinforcement of desired actions combined with sustainment of undesired reinforcement could help to promote a change.

Organisational Theories emphasis that organisation contexts also play a role in the process of change and they are also thought, like individuals, to pass through a sequence of stages. One model suggest three stages in the process of change: the "unfreezing" of old behaviours or practices which are no longer sustainable, "changing" to a new position by exposure to new information and a "refreezing" of the new position through reinforcement and support.

Other important issues to consider in this theory are the complexities of organisational change and the need to take account of context (*i.e.* why and when change should take place), process (*i.e.* how change will be secured), content (*i.e.* what change will be possible) and internal and external environments.

Therefore the identification of barriers to change and the development of strategies to overcome them are likely to be of fundamental importance in promoting the uptake of research findings by health professionals. It is essential to review the identified barriers during the process of implementation as their nature may change over time (Haines and Donald 1998).

Section 3: Methodological Consideration

2.3 Introduction

Evaluating the results from different published studies can be complex and many issues should be taken into consideration before these findings can be extrapolated to a particular population and employed for daily clinical decision-making (Sterne *et al* 2001). As many existing studies have used weak research designs or are methodologically flawed with potential threats to validity, it thus limits their value for informed decision-making (Bero *et al* 1998). Besides, the publication of a paper in a peer-reviewed journal is no guarantee that the design, methodology, analysis or the conclusion of an investigation is correct. Bias is the outcome of defects and flaws in the design or methodology of a study. Therefore this chapter seeks to describe briefly a number of potential study designs which are commonly used for evaluation of implementation strategies. Since the randomised controlled trial is the chosen design of this investigation, this chapter attempts to account for more detailed overviews of this design.

In addition, as the occurrence of methodological flaws and bias can jeopardise the validity of any investigation, a summary of the most common threats to validity and bias has also been included in this chapter.

The report of a trial, its aim, methodology and analysis should be conveyed to the reader in a transparent manner and a lack of adequately reported studies could be associated with bias in interpreting the results. Hence, a standard approach to reporting is vital in facilitating the improvement of the quality of the reports of trials. The CONSORT

(Consolidated Standard of Reporting Trials) statements for standards in the reporting of trials have also been published in recent years. The chapter will end with a brief account of the CONSORT history, its checklist items and flow diagram.

2.3.1 Study Designs

There are a variety of study designs which can be used for research in the evaluation of strategies for the implementation of change. A summary of these designs is presented as follows.

2.3.1.1 Observational studies

An observational study is the study of a single group of participants which can provide information about the process of behavioural change and generate hypotheses for further testing in rigorous evaluations (Grilli and Lomas 1994). However, they are rarely valuable for evaluating trials as the characteristics of the populations under comparison may differ in ways that affect the outcomes being measured. If these differences cannot be identified or measured, nothing can be done to revise the resulting bias. Even when it is possible to adjust for recognised differences, it is never possible to rule out unrecognised bias with confidence (Grimshaw *et al* 2000).

2.3.1.2 Quasi-experimental design

When a randomised controlled trial cannot be conducted due to practical and ethical issues, quasi-experimental studies are often undertaken. The three most common designs in implementation of change studies are I) uncontrolled before and after studies, II) time series designs and III) controlled before and after studies.

I) Uncontrolled before and after studies

Uncontrolled before and after studies measure participants' performance before and after the introduction of an intervention in the same study site. The change in outcome measured is assumed to be due to the intervention.

They are simple designs and superior to observational studies. However, they are inherently weak evaluative designs (Russell and Grimshaw 1992) as sudden changes or secular trends make it difficult to attribute the observed changes entirely to the intervention (Cook and Campbell 1979). The intervention is confounded by the Hawthorne effect which could lead to an overestimation of the effectiveness of an intervention. The "Hawthorne effect" can be defined as a change in participant performance that occurs simply by the awareness of external attention which can affect an intervention (Randall and Cebul 1991) or the non-specific beneficial effect on performance of simply taking part in research (Grimshaw *et al* 2000). Hence the results of such studies have to be interpreted with great caution and they are not suitable for research evaluating intervention strategies (Grimshaw *et al* 2000).

II) Time series design

When it is difficult to randomise or identify an appropriate control group (e.g. following the dissemination of a national guideline) this design is suitable because it tends to detect whether an intervention has had an effect greater than the underlying trend (Cook and Campbell 1979). Hence data are collected at series of times pre and post-intervention. The design increases the confidence in which the estimate of effect can be attributed to the intervention. However it does not provide protection against the effects

of other events occurring at the same time as the study intervention which might improve performance (Grimshaw *et al* 2000).

III) Controlled before and after studies

In this design, a control population is identified which has similar characteristics and performance to the study population and expected to experience similar secular trends and sudden changes to the study population (Cook and Campbell 1979). Data are collected in both populations simultaneously using similar methods before and after the intervention is introduced in the study population. An intra-group analysis comparing performances in the study and control groups following the intervention is undertaken and any observed differences are assumed to be due to intervention (Grimshaw *et al* 2000).

Although well designed before and after studies can protect against secular trends and sudden changes, it is often difficult to identify a comparable control group. Even in apparently well-matched control and study groups, performance at baseline is often different. Baseline imbalance (see Section 2.3.2) suggests that the control group is not comparable and may not experience the same secular trends or sudden changes as the intervention group. Thus the obvious effect of the intervention may not be valid (Cook and Campbell 1979, Grimshaw *et al* 2000). The usefulness of controlled before and after studies is also limited because the estimates of effect cannot be attributed to the intervention with confidence due to the non-randomised control group (Grimshaw *et al* 2000).

2.3.1.3 Balanced incomplete block design

In implementation research, there are a number of non-specific effects which may influence the estimate of the effect of an intervention such as the Hawthorne effect (see Section 2.3.1.2). If these effects are imbalanced across study groups in implementation of change trials, the resulting estimates of effect may be biased. The balanced incomplete block designs can be used to equalise such a non-specific effect and thereby minimised its impact (Cargill *et al* 1986). The simplest design is a 2X2 balanced incomplete block design which is the simultaneous arrangement of two trials in parallel. The study participants are allocated randomly between two groups. Group one receives the first intervention and provides control data for group two. Group two receives the second intervention and provides control data for group one.

The Hawthorne effect can be equalised across the two groups as participants in both groups experience the same level of intervention. Such design can enhance the generalisibility of the outcome results as they test the effects of the intervention across different conditions. However they are complex to design, conduct and analyse (Grimshaw *et al* 2000).

2.3.1.4 Randomised-controlled trials (RCT)

The randomised-controlled trial (RCT) defined as a quantitative, comparative, controlled study which investigates two or more interventions in a series of participants who receive them in random order (Health Technology Assessment 1999). RCT is seen by many as the most robust method of assessing health care innovations, especially those relating to therapeutic research in the last century (Prescott *et al* 1999). However, its history goes as far back as the eighteenth century when a Flemish physician, Van

Helmont, proposed a therapeutic trial of bloodletting for fevers and in 1747 ship's surgeon James Lind carried out trials of oranges and limes to treat scurvy. In the 1940's, the British Medical Research Council (MRC) in collaboration with drug licensing bodies tested streptomycin in a randomised-controlled trial. Austin Bradford Hill, as the statistician on the committee, had suggested that a statistical method of "randomisation" be used to determine which treatment group each patient should be placed into. It was Hill who promoted this as the first randomised-controlled trial to be trialed (Hill 1990). It is as a result of Hill's efforts that the streptomycin trial is believed to be the first RCT and that 1948 is celebrated as marking the beginning of a new era in modern medicine.

Similar developments occurred in other countries during the first half of the twentieth century and randomised-controlled trials became widely accepted as one of the simplest and most powerful tools in therapeutic research (Prescott *et al* 1999). Typically, randomised-controlled trials seek to quantify and compare different outcomes that are present or absent after participants randomly receive the interventions. Random allocation means that all participants have the same chance of being assigned to each of the intervention groups. Therefore, allocation is not determined by the investigators or the study participants (Campbell and Grimshaw 1998). The purpose of random allocation is that the characteristics of the participants are likely to be similar across groups at the start of the comparison. This helps the investigators to isolate and quantify the impact of the interventions that they are studying, with minimal effects from other factors that could influence the course of the participants. The most frequent unit of allocation in RCTs is the individual participant, either a patient or a clinician.

Sometimes, however, it is more appropriate to randomise groups of people rather than individuals. Examples of these groups or "clusters" are hospitals, practices, families and geographical areas. This is known as cluster randomisation (Campbell and Grimshaw 1998, Wilson *et al* 2000). This approach is taken when the RCTs are designed to evaluate interventions that may affect more than one individual within a particular group and the outcome for an individual can no longer be assumed to be independent. It is also used when the individual participant in one study group is likely to affect the assessment of other participants in the same groups. This phenomenon is known as contamination (Campbell and Grimshaw 1998, Wilson *et al* 2000). Contamination of control participants has two related effects. It reduces the point of estimate of an intervention's effectiveness and rejection of an effective intervention as ineffective because the observed effect size was neither statistically nor clinically significant (Torgerson 2001).

2.3.1.4.1 Advantages and disadvantages of randomised-controlled trials

The principal strength of randomised-controlled trials is that they minimise bias and have the ability to deal with known and unknown confounding factors (Cook and Campbell 1979, Torgerson 2001). When participants are randomly assigned to different intervention groups, the constitution of each group is generally similar and the randomisation controls the many threats to internal validity (see Section 2.3.4.1). Hence, in a RCT, threats such as selection, maturation or selection–maturation biases (explained later in Section 2.3.4.1) are absent. There are also no testing or instrumentation threats (will explain later in Section 2.3.4.1), since each participant experiences the same experimental conditions, measuring and observing instruments.

There is no deliberate selection of subjects since they are randomly assigned to the intervention and there is no statistical regression threat (explained later in Section 2.3.4.1). Each group experiences the same global pattern of history so there is no history bias in a RCT. If there are intervention-related differences in who drops out of the study, this is interpreted as a consequence of the intervention (Cook and Campbell 1979). Detection bias is avoided by arranging the outcome to be assessed unaware of the intervention groups. Co-intervention bias in a RCT is minimised by blinding interventions where possible and by employing clearly described intervention policies, which are identical for each group, apart from the intervention under examination (Prescott *et al* 1999).

Nevertheless the randomisation in RCT does not rule out all the threats even when the randomisation has been successfully implemented and maintained (Cook and Campbell 1979). In particular biases such as the imitation of interventions, compensatory rivalry and demoralisation (see Section 2.3.4.1) in groups receiving less desirable intervention or in non-intervention groups cannot be prevented by randomisation.

Furthermore, there are many situations in which RCT trials are not practical, necessarily appropriate or even sufficient to answer all clinical research problems and they can be ethically problematic at times. Randomisation may be inappropriate where a trial would have to be of disproportionate size and duration and thus costly, for example, if it were to evaluate the effect of intervention with a very rare outcome or with effects that take a long time to develop. It also may be inappropriate where the process of random allocation may affect the effectiveness of the intervention. This can arise when the

subjects cannot be blind to the intervention because the intervention requires their active participation, which in turn will be affected by their underlying beliefs and preferences. An example would be a trial of the effectiveness of clinical audit in improving the quality of patient care which would be complicated because the definition of effectiveness depends on the attitudes of participating clinicians (Black 1992).

Therefore with all of these factors in mind, to conduct a randomised controlled trial as accurately as possible meticulous attention has to be given to each stage of the trial to avoid any contaminations or bias.

2.3.1.4.2 Types of randomised controlled trials

There are many terms used to describe and classify randomised controlled trials and they can be categorised according to (Jadad 1998):

1. The different aspects of the interventions investigators want to explore
2. The way in which the participants are exposed to the interventions
3. The number of participants in the study
4. The presence, absence or degree of strategies to control bias
5. Whether the preferences of non-randomised individuals and participants are taken into account in the design of the study.

Each of these categorises are briefly described below (Jadad 1998), (see Table 2.3).

1 RCTs that explore different aspects of interventions

In this category RCTs can be classified as explanatory or pragmatic or as efficacy or effectiveness trials.

The explanatory trials address whether or not an intervention works and attempt to establish how such intervention works. They usually include a set of strict inclusion criteria which results in homogenous study groups. They tend to use placebos as controls.

Pragmatic trials or management trials are designed not only to determine whether the intervention works but also to describe all the consequences of its use, good and bad, under circumstances mimicking clinical practice (Sackett and Gent 1979). Pragmatic studies tend to include participants with heterogeneous characteristics, similar to those seen by clinicians in their daily practice, and use active controls.

RCTs are often described in terms of whether they evaluate efficacy or effectiveness in an intervention. Efficacy refers to whether an intervention works for the individual who receives it (Fletcher and Fletcher 1996). Efficacy trials tend to be explanatory trials, as they are designed to yield a 'clean' evaluation of the effects of the intervention. However they are not so interested in assessing how the intervention works. Their main goal is to include participants who will follow instructions and who will receive the intervention. Effectiveness refers to whether an intervention works for the individual to whom it has been offered. The effectiveness trials tend to be pragmatic, as they attempt to evaluate the effects of the intervention in a situation similar to those of daily clinical practice.

2 RCTs according to the participants' exposure to the interventions

Depending on the extent to which the participants are exposed to the study interventions, RCTs can have parallel, crossover or factorial designs.

In a RCT with a parallel group design, each group of participants is exposed to only one of the study interventions. For instance, in an evaluation of the effect of a new analgesia for toothache, the new analgesic would be given to one group of patients and a placebo to a different group of patients. Employing this design enables the investigators to carry out inter-group comparisons of the effect of this new analgesic with those of a placebo.

In a cross-over design RCT, each of the participants is given all the study interventions in successive periods. The order of given interventions to the participants is determined at random. This design produces intra-participant comparisons, as each participant acts under his or her own control.

An RCT has a factorial design when two or more experimental interventions are not only evaluated separately but also in combination (interactions between them) and against a control group.

3 RCTs according to the number of participants

RCTs can include one or more participants, they can have fixed or variable (sequential) numbers of participants and they can involve one or more centres.

'Individual patients' trial or 'n-of-1 trials' is basically a cross-over trial with one participant in which she/he receives the experimental and the control interventions, in pairs, on multiple occasions and in random order. These trials provide individual results which are not generalisable.

'Mega-trial' is a term that is used to describe RCTs with a simple and pragmatic design which includes thousands of patients, limited data collections and are multi-centre trials (Woods 1995). The main purpose of these trials is to gain 'increased statistical power' and achieve wider generalisibility.

A sequential trial is a study with parallel design in which the number of participants is not specified by the investigators in advance. The recruitment of participants continues until a clear benefit of one of the interventions is observed or the investigators are convinced that there are no differences between them.

In a fixed size trial, the number of participants or sample size is decided in advance. This can be done arbitrarily or calculated using statistical methods. The main advantage of using statistical methods to calculate sample size is to maximise the chance of detecting a statistically and clinically significant difference between interventions when a difference exists (Campbell *et al* 2001).

4 RCTs according to the presence, absence or degree of strategies to control bias

Depending on the extent of blinding, RCTs can be classified as open, single-blind, double-blind, triple-blind or quadruple-blind.

An open-blind RCT is a randomised trial in which everybody involved in the trial is aware of the allocations of the interventions to each participant.

A single-blind RCT is a randomised trial in which one group of individuals involved in the trial is unaware of the identity of the intervention that is allocated to each participant. Usually it is the participants or the investigators assessing the outcomes who are blinded to the identity of the interventions.

A double-blind RCT is a randomised trial in which two groups of individuals involved in the trial are unaware of the intervention that is allocated to each participant. Usually, these two groups include the participants and investigators in charge of assessing the outcomes of interventions.

Triple-blind and quadruple-blind are the same as double blind but with more groups of individuals unaware of the allocation of the intervention to each participant. These groups could include the participants, the investigator giving the intervention or those evaluating the outcomes and the data analysis.

5 RCTs that take into account the preferences of non-randomised individuals and participants

There are at least three types of preference RCTs that take into account the preferences of eligible individuals, as to whether they take part in the trial or not. Preference trials are rarely used in health care research and more likely to become more frequently used as consumer participation in health care decisions and research (Torgerson and Sibbald 1998). These trials can have Zelen's design, a comprehensive cohort design or Wennberg's design.

In a trial with Zelen's design, eligible individuals are randomised before they give consent to participate in the trial, to receive either standard treatment or an experimental intervention. Those who are allocated to standard treatment are given the standard treatment without knowing that they are part of a trial, whereas those allocated to the experimental intervention are offered the experimental intervention and told that they are part of a trial. If they refuse to participate in the trial, they are given the standard intervention but are analysed as if they had received the experimental intervention (Zelen 1979). To overcome the ethical concerns of not telling patients that they have been randomised to receive the standard treatment, the original design proposed by Zelen can be modified by informing participants of the group they have been allocated and by offering them the opportunity to switch to the other group. This design is known as double randomised consent design. This design is associated with the lack of blinding and the potential loss of statistical power.

A comprehensive cohort trial is a study in which all participants are followed up, regardless of their randomisation status. In these trials, if a person agrees to take part in an RCT, he or she is randomised to one of the study interventions. If a person does not agree to be randomised because of his or her strong preference for one of the interventions then that person will be given the preferred intervention and followed up as if he or she were part of a cohort study. At the end, the outcomes of this non-randomised group can be compared with those in a randomised group (Brewin and Bradley 1989).

In trial with Wennberg's design, eligible individuals are randomised to a 'preference group' or an 'RCT group'. Those individuals in the preference group are given the opportunity to receive the intervention that they choose, whereas those in the RCT group are allocated randomly to receive any of the study interventions, regardless of their preference. The outcomes associated with each of the interventions in each of the groups are compared and used to estimate the impact of the participants' preferences on the outcomes.

Table 2.3 Different types of RCTs

RCTs according to the aspects of the interventions they are to evaluate

- Explanatory and pragmatic trials
- Efficacy and effectiveness trials

RCTs according to how the participants are exposed to the interventions

- Parallel trials
- Cross-over trials
- Trials with factorial design

RCTs according to the number of participants

- From n-of-1 to mega-trials
- Fixed size
- Sequential trials

RCTs according to whether the investigators and participants know which intervention is being assessed

- Open trials
- Single-bind trials
- Double-bind trials
- Triple- and quadruple-bind trials

RCTs according to whether the preferences of non-randomised individuals and participants are taken into account

- Zelen' s design
- Comprehensive cohort design
- Wennberg's design

Jadad 1998

Now that the description of a variety of designs for the evaluation of implementation trials has been given, there are several other essential issues to be covered. They are "baseline equivalence" and the guidelines for the reporting of the randomised controlled trials "CONSORT". In addition, the methodological flaws and biases that can occur in different stages of a trial can jeopardise the validity of any investigations.

2.3.2 Baseline equivalence in randomised trials

In a randomised controlled trial, the participant characteristics that may influence outcome are distributed between groups under comparison so that any difference in the outcome can be assumed to be due to the intervention. Imbalance between groups in the baseline variables such as age or severity of a disease can bias the outcome and statistical tests. This property is sometimes referred to as chance bias. Observed differences in the outcome between groups could by chance be due to characteristics of the patients and not intervention. To avoid this chance bias randomisation and adjustment in statistical analysis for baseline variables could be the answer. Successful randomisation can result in balancing these confounding factors across the study and across control groups with a subsequent increase in statistical power (Grimshaw *et al* 2000). The baseline balance needs to be checked between the groups by statistical tests of baseline characteristics (Randall and Cebul 1991). These baseline measurements can then be used to assess the adequacy of the allocation process (Grimshaw *et al* 2000). Ideally such a measurement should be undertaken at pre-intervention, in the planning or pilot stage of an implementation study. These measurements can be useful for providing an estimate of the magnitude of a problem. A low compliance score at pre-intervention may indicate there is much room for enhancement. On the other hand, high

compliance scores may indicate that there is little opportunity for improvement (ceiling effect), (Grimshaw *et al* 2000).

2.3.3 CONSORT Statements

The problem that most readers of RCT reports often encounter is that these reports lack many features that are useful in evaluating their quality. Readers must understand the designs, methodology, analysis and interpretations of the reports to comprehend their results and conclusions. To achieve this, the authors should convey their research to the reader by transparent and clear reporting.

To improve the quality of reports of RCTs, in the mid 1990s, an international group of clinical researchers, statisticians, epidemiologists and biomedical editors developed and published the Consolidated Standards of Reporting Trials (CONSORT) statement (Moher *et al* 1994, Needleman 2000, Moher *et al* 2001, I). The CONSORT statement is a checklist of items and a flow diagram that aims to help authors improve the quality of reports of RCTs. However, the basic idea underlying the development of the statement can be applied to any design. The checklist items require clear and thorough descriptions of hypothesis, protocol, randomisation, allocation concealment, follow up and analysis.

The CONSORT statement is continuously evolving with time and CONSORT group members meet regularly to review the need to refine the CONSORT statement. In 1999, the merits of including each item in the light of current evidence were discussed and the original statement was revised (Moher *et al* 2001,I). Early results indicate that the use of

CONSORT has helped to improve the quality of the reports (Moher *et al* 2001, II). In 1994, in a review of 71 published RCT reports in 3 journals, allocation concealment was not reported clearly in 43 (61%) of the studies. Four years later after these journals required that authors use the CONSORT statement for reporting the RCTs, the proportion of articles in which allocation concealment was not clearly reported had decreased to 30 out of 77 (39%), (mean difference: -22%; [CI 95%, -38% to –6%]) (Moher *et al* 2001, II). Therefore the use of the CONSORT statement reduces the number of inadequate reporting of RCTs and potentially positively influences the manner in which RCTs are conducted (Moher *et al* 2001, I).

2.3.4 "Bias" and "Validity" of a study

When appraising the findings of a research study it is imperative to assess the potential influence of each source of bias which may threaten its validity. Bias can be defined as "any process at any stage of a research study which tends to produce results or conclusions that differ systematically from the truth" (Murphy 1976). It is distinct from random error which occurs by chance. There are many different types of bias that can occur at different stages of a study (Table 2.4) (Sackett 1979).

Table 2.4 Different stages of research in which bias can occur

Stages of research are

- In reading-up on the field
- In specifying and selecting the study sample
- In executing the experimental manoeuvre or exposure
- In measuring exposure and outcome
- In analysing the data
- In interpreting the analysis
- In publishing the results

Sackett 1979

The validity of a study may be defined as the degree to which the conclusions drawn are correct. The two main types of validity in relation to a study design are internal validity and external validity.

Internal validity is defined by Cook and Campbell as " the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of cause" (Cook and Campbell 1979). They also defined external validity as "the approximate validity with which we can infer that the presumed causal relationship can be generalised to and across alternate measures of the cause and effect and across different types of persons, settings, and times" (Cook and Campbell 1979), in other words, the degree of generalisibility to other settings.

Cook and Campbell (1979) explained, "estimating the internal validity of a relationship is a deductive process in which the investigator has to systematically think through how each of the internal validity threats may have influenced the data". Then, the investigator has to examine the data to test which relevant threats can be ruled out.

Throughout this process, the researcher has to be his or her own severest critic and rigorously examine all of the threats the researcher can imagine. When all of the threats can be plausibly eliminated, it is possible to make confident conclusions about whether a relationship is probably causal. To assist with the deductive process of evaluation of the study validity, they provided an index of common threats to the different types of validity (Cook and Campbell 1979) and a summary of these is outlined below.

2.3.4.1 Threats to Internal validity

Cook and Campbell provided an inventory of specific threats to internal validity which are outlined below (Cook and Campbell 1979):

- History: External events, which take place between the pre- and post-intervention measurement may influence the observed effect over and above the intervention effect.

- Maturation: The passage of time may bring changes in the subjects of a research independent of the intervention, such as deterioration or improvement of their performance.

- Testing: Administering a pre-test or baseline measurement may alter the response to later measurement.

- Instrumentation measures: Measurement bias occurs when the measurement of exposure and/or outcome are not valid or the measuring instrument used to evaluate outcomes may change over time and the effect recorded may not be due to intervention alone. This bias can be introduced by the observer (observer bias), by the participant (responder bias) or by the instruments (instrument bias) used to make the measurements and record the data such as a questionnaire. This bias may be minimised or eliminated by a standardisation of measurement process (Moles and dos Santos Silva 2000).

- Regression to the mean: Experimental units selected on the basis of their extreme scores will tend to give subsequent scores closer to the average.

- Selection: This a threat that may occur when the outcomes of a trial are affected by systematic differences in the way in which subjects are selected for inclusion in a study and are allocated to groups under comparison. This threat can often occur with self-selected and volunteer groups as they tend to be different from the general population.

- Mortality (differential attrition): An effect that is due to systematic differences between subjects under comparison, *i.e.* those who withdrew from the investigation and those who remain in the study during the course of the experiment. This results in a selection artefact. Theoretically this can be avoided

by scrupulous tracking of those withdrawing from the trial, accompanied by an explicit description of protocol deviations (Schulz *et al* 1995).

- Selection maturation interaction: This occurs when time-dependent changes vary systematically in different types of experimental units. Selection maturation results when subjects are maturing in different speeds. Selection history results from various intervention groups coming from different settings so different groups could experience a unique local history that might affect outcome variables. Selection–instrumentation occurs when different groups score at different mean positions on a test whose interval is not equal.

- Diffusion or imitation of treatment/performance bias is systematic differences in intervention provided to groups under comparison other than the target intervention (Schulz *et al* 1995). It commonly takes one of two forms: contamination (application or diffusion or imitation of the intervention intended only for the intervention group to some portion of the controls) or co-intervention (provision of unintended intervention to either arm of the trial). When contamination between intervention and non-intervention groups occurs, these effects may lead to an effective intervention appearing ineffective.

- Compensatory rivalry: When performance in the non-intervention group changes by becoming motivated. This can reduce or reverse the expected difference between the non-intervention group and the others.

- Resentful demoralisation: When performance in the non-intervention group deteriorates following the introduction of an intervention to the other study group. This can result in an ineffective intervention appearing effective.

Further statistical conclusion validity was considered, which is a part of internal validity. Its evaluation is imperative since the threats to statistical conclusion validity may result in the probability of type I error (false positive) and type II error (false negative) in a study. Statistical conclusion validity is concerned with sources of random error and with the appropriate use of statistical tests. The list of these threats is outlined as follows:

- Low statistical power: The chances of making type II error (no-difference conclusion) increases when sample sizes are small and the probability level (α) is set low.

- Violated assumption of statistical tests: Most statistical tests of null hypothesis require certain assumptions to be met and if these are violated there is the possibility of making incorrect interpretations of the results of the data analysis.

- Fishing and the error rate problem: The likelihood of type I error increases when multiple comparisons tests are made.

- The reliability of measures: There is a possibility of inflating error variance (Type I or Type II error) if measures of outcome of interest have low reliability. There are some ways of increasing the reliability of measures such as using longer tests for which subjects or measures have been chosen for their great inter-correlations or using groups (or aggregated units) rather than individuals due to the higher stability of group mean.

- The reliability of treatment implementation: The lack of standardisation and homogeneity in delivery of intervention across subjects and occasions may result in decreasing the likelihood of obtaining true results (Type II error).

- Random irrelevancies in experimental setting: There is the possibility of inflating error variance (Type II error) if some features of the experimental setting other than treatment or intervention affect the outcome measure.

- Random Heterogeneity of respondents: The error variance (Type II error) will be inflated if there is a great heterogeneity of participants.

2.3.4.2 Threat to external validity

The external validity concern is with generalising the finding of a study beyond the sample population to various populations, settings and times. However, the problem of systematic recruitment can lead to findings that are only applicable to a selected group such as volunteers. The volunteer will often have significantly different characteristics

from the target group in whom the practice variation may be most marked. They will be the most progressive, overconfident and institutionally exhibitionist (Cook and Campbell 1979, Grol and Jones 2000). One way of reducing this bias is to make collaboration in the research as convenient as possible for the subjects.

Cook and Campbell proposed these three models for possibly increasing external validity:

Random sampling for representativeness model: The most representative sample is randomly chosen from the population and then randomly selected units are randomly assigned to various intervention groups. However, this model requires considerable resources since it can involve a large scale and multistage sampling area. It is feasible though when sampling a targeted population of people rather than settings or historical times. Hence this model is a powerful model for generalisation of findings where the targets are specified. The important function of random samples is to permit examination of the data for differential effects on a variety of subpopulations.

Model for heterogeneity: a target class of persons, settings and times is defined and deliberate purposeful sampling is carried out for heterogeneity. Sampling design in this model does not involve random sampling, hence it is doubtful if one can generalise the findings from the achieved samples to any other populations. However, a wider range of heterogeneous population from different backgrounds, settings and time has been selected rather than samples from a homogenous group (e.g.: sampling school children from cities, towns and rural settings and to test whether an intervention has comparable effects in each of the subgroups of the children and settings).

Impressionistic model instance model: it is feasible when sampling an impressionistically very similar sample to an intended population, setting or time to one which wants to generalise the findings. It is the least powerful model. However, it can be used with small numbers of samples for convenience.

Construct validity, which is a part of external validity, refers to the possibility that a particular underlying relationship can be made by more than one construct relationship (for example: A can cause B, however, it is possible that A caused X and X caused B). Construct validity is concerned with "confounding" factors where the intervention outcome may be affected by other underlying factors beside the intervention (Moles and dos Santos Silva 2000). If it is possible, the construct and confounder factors should be carefully clarified prior to the initiation of the experiment in order for definitions to be clear. Subsequent data analysis should be conducted in a way that the proposed dependent variables should control for relevant confounding factors and variables that are inadvertently manipulated at the same time as the intended intervention or that are inadvertently measured as part of an effect construct (Cook and Campbell 1979).

Cook and Campbell (1979) suggested an inventory of some of the relevant threats to the construct validity of assumed cause and effects which are listed below:

- **Inadequate preoperational clarification of construct:** A precise clarification of cause and effect construct is vital for a high construct validity since it permits tailoring of the statistical tests to whichever definitions emerge from the clarification.

- **Mono-operational bias:** Construct validity will be lower in single intervention research than research involving multiple interventions (e.g. factorial design). The use of single intervention research under-represent constructs since there is one example of cause and only one measure to represent each possible effect construct. In multiple interventions, the variation due to the difference between interventions can be examined to test whether the different combinations of confounder factors will affect the outcome differently and whether each intervention singly caused the expected outcome. Besides it is not as costly to gather additional data in factorial designs research.

- **Mono-method bias:** A particular method used for measuring or collecting data itself can be confounders in a study and its influence cannot be dissociated from the influence of the main construct of the study (e.g. when collecting data by questionnaires as opposed to using other methods. It would be more accurate to say the outcome was presented from a questionnaire survey and it would not be possible to generalise about other situations in which for example, data are collected by interviewing).

- **Hypothesis-guessing within experimental conditions:** is a situation where participants try to guess what experimenters are hoping for or expecting them to behave. This can be best avoided by making the hypothesis hard to guess by decreasing the general level of reactivity in the experiment.

- **Evaluation apprehension:** When the participants are apprehensive about being evaluated by experts in the field.

- **Experimenter expectancies:** Human behaviour can be influenced by previous experiences, knowledge, expectation or belief, and in research particularly, there is the chance of expectation influencing findings. This is important especially when some subjectivity exists in assessment, leading to a biased outcome (Day and Altman 2000). Bias in decision-making can occur when using expert opinion, personal clinical experience or published research.

Ultimately, there are trade-offs between one type of validity and another and some of them are unavoidable. For research which is concerned with the effectiveness of an intervention the priority ordering begins with internal validity then the external validity, since the magnitude and direction of causal relations and tests of a causal hypothesis is more important than a desire for the generalisation of the results to other settings (Cook and Campbell 1979).

As far as bias is concerned, the goal of any clinical trial is to minimise or even eliminate all bias. However when bias cannot be eliminated, it is imperative to define any sources of the error so that they may be taken into consideration when conclusions are drawn from any research findings.

Section 4: Third Molar Teeth

2.4 Introduction

As the treatment of third molar teeth has been a subject of debate for decades and with the recent SIGN guideline on "management of unerupted and impacted third molar teeth" (SIGN 43, 2000) being used to evaluate different interventions employed in this trial, an account of the different aspects of these teeth will be given in this section.

2.4.1 Background

Although a wide variation exists in the eruption dates of third molar teeth, these teeth generally erupt between the ages of 18 and 24. One or more third molars are absent in approximately 25% of adults (Song *et al* 1997, Von Wowern *et al* 1998) but they may still be present in elderly, otherwise edentulous patients.

The term unerupted refers to a tooth lying within the jaws but entirely covered by soft tissue and partially or completely covered by bone. Partially erupted means a tooth that has failed to erupt fully into its normal functional position (The Royal College of Surgeons of England Faculty of Dental Surgery 1997). The term implies that the tooth is partly visible or in communication with the oral cavity. An impacted tooth is a tooth which is prevented from completely erupting into a normal functional position. This may be due to lack of space in the dental arch, obstruction by another tooth or bone or an abnormal eruption path (SIGN 43, 2000). It should be considered that normally erupted teeth all used to be partially erupted or unerupted at some stage of their eruption pathway. Hence, partially erupted or unerupted teeth may not be impacted (The Royal College of surgeons of England Faculty of Dental Surgery 1997). Nevertheless the

failure of eruption of third molars in particular is a very common condition (SIGN 43, 2000).

2.4.2 Epidemiology of impacted third molars

The prevalence of third molar impaction varies widely and is influenced by age, gender and ethnicity. It is also dependent on the definition of impaction and the characteristics of the population studied. The estimated prevalence lies within the range of 8% to 66% (Schersten *et al* 1989, Ahlqwist and Grondahl 1991, Robinson 1994). Impacted or unerupted teeth are not, in themselves, pathological (Brickley *et al* 1995, I) but the impaction may increase the risk of disease, particularly when oral hygiene is poor (SIGN 43, 2000). The pathology which may develop in association with an impacted third molar, such as pericoronitis, can cause painful symptoms and can have a significant impact on the individuals involved. However, the prevalence and incidence of third molar pathology among patients with impacted third molar has rarely been reported in literature.

The pathologies associated with impacted third molars and available epidemiological data is outlined below:

2.4.2a Pericoronitis:

Pericoronitis is defined as inflammation of the gingival tissue surrounding the crown of a tooth and is the most common reason for the removal of impacted third molars (Worrall *et al* 1998). Its prevalence associated with impacted third molars has not been widely studied. A 10% incidence of pericoronitis has been reported in a study by Von Wowern and Nielsen (1989) in which 130 lower third molars of 70 students were kept

under observation for a period of 4 years. This is in agreement with the report from Worrell *et al* (1998).

Von Wowern and Nielsen (1989) chose students as their subjects, who might have better oral hygiene than the rest of the population.

2.4.2b Cysts and Tumours:

Song *et al* (1997) reported the incidence of cyst formation involving impacted teeth in their systematic review from 0% to 11%. These estimates may not be reliable since the details related to the incidence were not often presented (Song *et al* 1997). Stephen *et al* (1989) suggested that the incidence of dentigerous cysts in many studies had been exaggerated by confusion with enlarged follicular space and by not having histopathological confirmation.

Daley (1996) reported that cyst development is very rare and that the risk of malignant neoplasm arising in the dental follicle is negligible and is not considered to be an indication for prophylactic removal.

Guven *et al* (2000), in a retrospective study of 9994 impacted third molars in referral patients to a dental hospital setting, demonstrated that the incidence of cyst formation associated with impacted third molars was 2.31% and that impacted third molars had a low incidence (0.8%) of tumour formation. These figures could be biased due to the number of referrals (41% of patients were referred by GDPs and 43% from the oral diagnosis department).

2.4.2c Caries and periodontal problems:

Mercier and Precious (1992) in their review quoted the incidence of periodontitis associated with impacted third molars as ranging from 1% to 4.5% (they did not report the assessment of the validity of primary studies).

Very few impacted third molars cause dental caries of second molars (Daley 1996) and this study reported a prevalence of 1% for caries (but his review lacks information about review methods). Brokaw (1991) and Tate (1994) suggested that caries problems are common but that there was no objective evidence for the finding of caries associated with impacted third molar teeth (Song *et al* 1997 *op. cit.*).

Van der Linden *et al* 1995, in a review of 1001 patients aged 13 to 75 whose third molars were removed reported the prevalence of caries in 7% (204/2872 teeth) of impacted third molars and in about 5% (1227/2872 teeth) of adjacent molar teeth.

2.4.2d Resorption of second molars:

Two studies by Brokaw (1991) and Tate (1994) suggested that impacted third molars frequently caused resorption of second molars but no objective evidence was provided. Nitzen *et al* (1981) reported a 7% incidence of root resorption of second molars adjacent to an impacted third molar tooth.

Mercier and Precious (1992) in their review quoted the prevalence of root resorption of second molars ranging between 0% - 3.1%.

Also a low incidence of less than 1% of root resorption of second molars with impacted third molars was suggested by Von-Wowern and Neilson (1989). In a review by Daley (1996), it was concluded that the risk of second molar root resorption by an impacted third molar is low and is likely to occur in younger patients for whom surgery is claimed to be associated with lower morbidity.

2.4.2e Third molars and crowding:

Song *et al* (1997) stated in their review that the association between crowding and impacted third molar teeth is not sufficiently significant to warrant removal of the third molar teeth for the prevention of anterior incisor crowding and this is in agreement with the report from Harradine *et al* (1998).

2.4.3 Indications for removal of mandibular third molar teeth

In principle, third molar surgery should result in a health gain either physical, psychological or socially. However, case morbidity, surgical complications, cost effectiveness and benefit should also be considered in the formation of a well-balanced management plan for third molar pathology.

Third molar surgery is not risk-free and complications and suffering following such surgery may be considerable (Mercier and Precious 1992). Therefore, several questions should be asked before initiating such treatment.

- Does the third molar tooth need to be removed? In other words, has the correct diagnosis been made? Is it necessary to remove the third molar tooth which may erupt successfully and have a functional role in the dentition?

- Is there any absolute medical contra-indication for surgical extraction?

- Is simultaneous removal of asymptomatic contra-lateral teeth necessary?

- Should deeply impacted third molars with no history or evidence of pertinent local or systematic pathology be removed?

- Have we considered the risk of surgical complications?

- Is the dentist capable of extracting the third molar tooth or should the case be referred to someone more expert in the field?

These questions should be considered in light of the existing evidence providing strong indications for the surgical removal of third molar tooth.

2.4.3a Strong indications for removal of third molars

Based on the Third Molar SIGN Guideline (43, 2000) strong indications for the removal of third molars are:

- π Removal of any symptomatic wisdom tooth, especially where there have been one or more episodes of infection such as pericoronitis, cellulitis, abscess formation or untreatable pulpal/ periapical pathology.

- π Removal of an unsalvageable carious third molar or removal when access is needed to treat caries in the adjacent second molar tooth.

- π Removal of third molar tooth in the case of periodontal disease due to its position and its association with the second molar tooth.

- π Removal of third molar in the case of a dentigerous cyst or other related oral pathology.

- π Removal of third molar in the case of external resorption of this tooth or the second molar where this would appear to be caused by the third molar.

Other indications are:

- π Removal of third molar may be indicated prior to orthognathic surgery or for autogenous transplantation to a first molar socket.

- π Removal may be considered in cases of fracture of the mandible in the third molar region or for a tooth involved in tumour resection.

- π Removal of an unerupted third molar in an atrophic mandible may be appropriate and the situation needs to be carefully evaluated.

- π Prophylactic removal of a partially erupted third molar or a third molar that is likely to erupt, may be appropriate in the presence of certain specific medical conditions.

- π Removal of partially erupted or unerupted third molars, close to the alveolar surface, should be considered prior to denture construction or close to a planned implant.

- π Acute exacerbation of symptoms occurring while the patient is on a waiting list for surgery may be managed by extraction of the opposing maxillary third molar. (SIGN 43, 2000)

2.4.4 Removal of lower third molars: complications

The common complications and risks following third molar surgery are sensory nerve damage, alveolar osteitis (dry socket), infection, haemorrhage, pain and difficulty in eating. Conditions such as severe trismus, oro-antral fistula, buccal fat herniations, iatrogenic damage to adjacent second molar and iatrogenic mandibular fracture constitute rare complications.

The common complications are outlined below:

2.4.4a *Inferior alveolar nerve injuries*

The reported incidence of inferior alveolar nerve injuries following mandibular third molar surgery varies in the literature between 0.4% and 7.8% (Bruce *et al* 1980, Kipp *et al* 1980, Rood 1983, Osborne *et al* 1985, Sisk *et al* 1986, Alling 1986, Rood 1992, Chiapasco *et al* 1993).

Rood (1992) reported the incidence of 7.6% labial sensory disturbance and 0.25% permanent injury with the lingual split technique and there was a similar figure of labial paraesthesia from the use of a drill with 2.2% permanent injuries.

2.4.4b *Lingual nerve damage*

The reported incidence of lingual nerve damage ranges from 0.6% to 23% (Rud 1970, Goldberg *et al* 1985, Mason 1988, Von Arx and Simpson 1989, Carmichael and McGowan 1992, Absi and Shepherd 1993). Rud (1970) reported a 1% rate in patients treated by the lingual split bone technique and no instances of nerve injury when a buccal approach was used. This was in agreement with the findings of Chiapasco *et al* (1993). Carmichael and McGowan (1992) reported a significant increase in the incidence of sensory deficit if a lingual retractor was inserted with no significant difference caused by the use of a drill or chisel. Rood (1992) reported that there is an appreciable incidence (13%) of temporary lingual nerve disturbance after the removal of third molars using the lingual split technique with chisels. There were no cases of permanent damage. When a surgical drill was used, there were fewer cases (3.2%) of temporary disturbance, but 2% of cases with permanent injuries. In this study, the

insertion of the Howarth periosteal elevator as a lingual retractor was common to both techniques, whereas permanent injuries seemed to be associated only with the use of the drill. It was thought that Howarth periosteal elevator was unlikely to be associated with permanent nerve injury although it did not appear to be a satisfactory protector of the lingual nerve. Similarly, a study of 771 patients by Robinson and Smith (1996) confirmed that the avoidance of lingual retraction reduces the incidence of temporary lingual disturbance and does not increase the incidence of permanent damage. This indicated that the use of the Howarth's in this way is invalid and suggested that if at all possible use of lingual retraction should be avoided.

In a recent study of 946 patients in a dental school clinic by Valmaseda-Castellon and his colleagues (2000), they reported risks of 2% temporary nerve damage and 0% for permanent nerve damage. Other factors such as tooth position, retraction of the lingual flap, prolonged operating time and surgical experience appeared to play more of a role in increasing the risk of nerve damage.

2.4.4c *Mylohyoid nerve damage*

Mylohyoid nerve damage gives an area of altered sensation at the point of the chin but most investigators do not ask specifically about this nerve. Therefore it is rarely mentioned in the literature. However Carmichael and McGowan (1992) found a low incidence of long-term damage of 0.07% of operated sites.

The likelihood of nerve damage in general during the removal of third molars under general anaesthesia is five times higher than under local anaesthesia. Brann *et al* (1999)

reported an incidence of 8% and 12% of cases with loss of sensation in the inferior alveolar and lingual nerve distributions respectively for the patient treated under general anaesthesia. There were 0% and 3% of cases with loss of sensation in the inferior alveolar and lingual distributions, respectively, for patients treated under local anaesthesia.

2.4.4d *Alveolar osteitis*

The reported incidence of alveolar osteitis or dry socket varies between 0% and 35%. (Mercier and Precious 1992, Chiapasco *et al* 1993). The risk of dry socket increases with lack of operator surgical experience and patient tobacco use (Larsen 1992).

2.4.4e *Postoperative Haemorrhage*

Incidence of excessive haemorrhage/bleeding following third molar extraction ranges from 0.6% to 5.8% (Bruce *et al* 1980, Osborne *et al* 1985, Sisk *et al* 1986, Stanley *et al* 1988, Alling *et al* 1993, Chiapasco *et al* 1993).

2.4.4f *Damage to adjacent teeth*

The reported incidence of damage to adjacent teeth is reported as 0.3%-0.4% (Sisk *et al* 1986, Chiapasco *et al* 1993).

2.4.5 Clinical requirement and current practice

The management of unerupted and impacted third molars has received considerable attention in recent years. This is because removal of third molar teeth is a most common practice (Shepherd 1993, Effectiveness Matters 1998, Eklund and Pittman 2001) frequently associated with varying degrees of morbidity (Robinson and Smith 1996, Ogden *et al* 1998, Song *et al* 1997) including pain, swelling and trismus together with the possibility of temporary or permanent nerve damage, resulting in an altered sensation in the lip or tongue and subsequent medicolegal and economic implications for patient and clinicians alike.

The removal of third molar teeth constitutes one of the highest volume surgical procedures within the UK (Shepherd 1993, Effectiveness Matters 1998). These interventions account for about 1/3-2/3 of all surgical interventions, which are about 25,000 annually in Sweden and 2.25 million in the US. These figures do not include the teeth that are removed by general dental practitioners (Knutsson *et al* 1992, I, op. cit). In 1994-95 there were over 36,000 in-patient and 60,000 day case admissions for surgical removal of these teeth in England (Effectiveness Matters 1998, op. cit). Approximately 90% of patients on waiting lists for oral and maxillofacial surgery are scheduled for the removal of third molar teeth (Shepherd *et al* 1994).

Current practice includes both the removal of impacted third molars causing pathological changes as well as the early prophylactic removal of pathology-free impacted third molars. There appears to be a substantial variation in the management of

patients with third molar problems (Gilthorpe *et al* 1997, Landes 1998, Eklund and Pittman 2000, Eklund and Pittman 2001) as detailed later in Section 2.4.1.7.

There is a consensus among dental clinicians regarding the criteria for the removal of third molars associated with local disease. However, the appropriateness of prophylactic removal of asymptomatic third molars has been debated for many years. A recent review by the NHS Centre for Reviews and Dissemination undertook a study of 12 systematic reviews and concluded, "in the absence of good evidence to support prophylactic removal, there appears to be little justification for the removal of pathology-free impacted third molars" (Effectiveness Matters 1998).

Besides, surgical removal of third molar teeth consumes substantial resources within the NHS annually. It has been estimated that the cost to the NHS in England adds up to £30 million per year (Landes 1998) with approximately £20 million spent annually in the private sector (Shepherd 1994).

2.4.6 Prophylactic removal of third molar teeth

The influence of preventive dentistry in the early seventies has led to a focus on the removal of asymptomatic pathology-free third molars on the basis that they may become associated with pathology at some time in the future (Laskin 1971), when a great risk of morbidity associated with surgical intervention would be a real concern in older age (Robinson 1994, Brickley *et al* 1996, I, op. cit.). This emphasis influenced practice for many years. However, pericoronitis, the most frequent disease related to retained lower third molar teeth (Brickley *et al* 1995, II), generally affects mostly young

adults aged 18-25 and the incidence of this disease decreases very sharply beyond the age of 30 years. There is only a small risk that an impacted tooth found in middle-aged and older individuals will become associated with a serious pathological condition even over a relatively long period of time. Ahlqwist and Grondahls' (1991) findings on Swedish females aged 38 – 60 over a 12-year period corroborates this claim, and this is also supported by work of Venta *et al* (2000) whose study of 81 university students in Finland over a period of 12 years indicated that the need for surgical removal of third molars decreases significantly from the age of 20 to 32 years. These authors did not recommend the routine prophylactic extraction of asymptomatic third molars in young adults (Venta *et al* 2000). A recent retrospective histopathological study of more than 2,600 pericoronal lesions of extracted partially erupted teeth showed a relationship between pathologically significant disease and age (Curran *et al* 2002). The extent of the risk of significant pathological changes in older adults is reported to vary between 0 – 12% (Effectiveness Matters 1998).

Despite the small risk of pathology in third molar teeth later in life, there are still substantial variations in the way dental clinicians judge the need for removal of asymptomatic, pathology-free third molars and make treatment decisions (Knutsson *et al* 1992 I and II; Kostopoulou *et al* 1997, 1998 I and II, 2000). There is a lack of evidence to support the prophylactic removal of impacted third molars and there appears to be little justification for the routine removal of pathology-free impacted third molars (Mercier and Precious 1992, Robinson 1994, Brickely *et al* 1995, Song *et al* 1997, Effectiveness Matters 1998, Venta *et al* 2000).

The number of third molar teeth currently removed for prophylactic reasons in oral surgery units in the United Kingdom is unclear. Brickley *et al* (1993) in a study conducted in Cardiff suggested that one in three third molar teeth were removed without apparently clear reasons for surgery. A subsequent study conducted in Cardiff by Brickley and Shepherd (1996) demonstrated that only 59% of third molar teeth scheduled for surgery had valid indications and Brickley *et al* (1996, II) suggested that 30%-60% of removal of third molars is inappropriate.

Lopes *et al* (1995) in a prospective investigation of 522 patients undergoing third molar surgery in a large London teaching hospital showed that over half of all third molars removal (50.7%) had been undertaken without any apparent clinical indications for intervention. In the same study they found that the patients who did not have clinically sound justification for surgery have a similar incidence of sensory deficit and morbidity (10.4%) when compared to those with valid indications.

In stark contrast to the evidence presented by Brickley *et al* (1993) and by Lopes *et al* (1995), in a retrospective audit of 454 consecutive patients referred for third molar surgery to a maxillofacial unit of a large district general hospital trust, Pratt *et al* (1998) demonstrated that 97% of patients had at least one valid indication for surgery. Pratt *et al* explained this difference by stating the lack of accepted, clear guidelines within the department might have resulted in the formulation of haphazard treatment decisions. Unguided or inexperienced assessors may simply opt to agree with the referring practitioner that extraction is warranted. Two studies by Brickley *et al* (1993) and Pratt *et al* (1998) specified that only experienced clinicians were involved in the initial

assessment but Lope *et al* (1995) did not identify the seniority of the clinicians who made the treatment decision. Neither study stated whether the assessing clinicians had used accepted departmental guidelines to formulate their treatment decisions apart from Pratt *et al* (1998).

The data giving rise to most of the assumptions in the above studies are of necessity derived from specific and limited populations and therefore a strong indication for removal of an impacted third molar should always be complimented with a strong contra-indication to its retention (Precious 2000).

2.4.7 Variation in provision of care

There is a substantial variation in the provision of care and there is a wide variation in the rates of surgery for third molar teeth internationally.

Eklund and Pittman (2000) investigated the patterns of third molar removal in private dental practices in Michigan in order to determine if there is a consistent standard of care. A total of 52,193 patients born between 1973-1977 were followed from 1990 to 1997. He found a substantial variation among primary care dentists. Practices ranged from no adolescent patients having third molars removed, to virtually all having all four third molars extracted. He concluded that in practice there is no widespread agreement concerning third molar removal.

When the removal of third molars was examined against the socio-economic background of patients, in a 5-year study, Gilthorpe *et al* (1997) found that there was

inequality of oral surgery provision to the deprived population. These communities are less likely to have their third molars removed than the more affluent population. In a study of the relationship between dental health and the level of third molar removals experienced by a population, Landes (1998) concluded that a quarter of the variations in the provision of third molar surgery are related to differences in dental health. Populations with poor dental health experience fewer third molar removals than populations with good dental health.

However, Landes (1998) stated that the reasons for these results were complex and might have been related to local variations in the referral behaviour of general dental practitioners. Alternatively it may be that young adults in the most deprived areas hardly ever visit a dentist, and even if they do, they are more likely to refuse referral for an asymptomatic condition. The decrease might even be explained by a general deterioration in dental health amongst deprived communities, with more extractions of the first and second molars and therefore less need for extractions of third molars. Nevertheless, the relationships demonstrated in these studies (Gilthorpe *et al* 1997, Landes 1998) are not strong enough to explain a wide variation in the levels of third molar surgery in different areas in Britain.

Furthermore, several studies (Knutsson *et al* 1992 I and II) have concluded that there is a wide variation between clinicians in terms of treatment decisions regarding pathology-free third molars, suggesting that these decisions are not being made using consistent criteria.

In two continuous studies undertaken by Knutsson *et al* (1992 I and II), 30 general dental practitioners and 10 oral surgeons were asked to evaluate the need for removal of asymptomatic mandibular third molars based on a radiograph and the short history of 72 cases. The authors reported that there was great variation among the general practitioners in their perception of the need for removal of asymptomatic mandibular third molars. The same degree of variation in judgement was also observed among oral surgeons. Oral surgeons proposed 33% of the third molars for extraction, about the same proportion as general dental practitioners. The range of judgements by oral surgeons varied from 8% to 58%, with those having longer experience proposing fewer extractions (Knutsson *et al* 1992 II).

Ten years later Knutsson and colleagues (2001 I) assessed the same general dental practitioners and oral surgeons by sending the same cases in order to examine their decisions on the prophylactic removal of impacted mandibular third molars. They found that the intra-individual agreement between the treatment decisions on the two occasions, estimated for each dentist's decision on each case, varied between 56% and 97% and there was no significant difference in the number of teeth designated for removal between the two occasions for both categories. They concluded that presented variations among dental clinicians can have a profound impact on patients' oral health and it appeared that there has been no change over ten years towards a more non-interventionist attitude despite the current evidence on the inappropriateness of the prophylactic removal of third molar teeth (Knutsson *et al* 2001 I).

In a comparison study of general dental practitioners and oral surgeons in two countries (Sweden and UK) as to decisions regarding the prophylactic removal of mandibular third molars by the same researchers (Knutsson *et al* 2001 II), they demonstrated that the decision to remove mandibular third molars is based more upon individual clinician factors rather than on the category of practitioner or country of origin/practice.

In this study clinical and radiographic information relating to a stratified sample of 36 disease free mandibular third molars was presented to 26 GDPs and 10 oral surgeons in Sweden and 18 GDPs and 10 oral surgeons in Wales. They were asked to decide whether or not each third molar should be removed. Their result demonstrated that there was no evidence of any difference in the mean number of molars scheduled for removal by the 2 groups of GDPs but the Swedish oral surgeons scheduled significantly more third molars for removal than oral surgeons in Wales. They concluded that the less interventionist approach among oral surgeons in the UK may reflect the development and application of authoritative guidelines in the UK and an extensive debate concerning appropriateness of prophylactic removal there (Knutsson *et al* 2001 II).

Kostopoulou *et al* (2000) suggested that the reason for variation in the treatment of third molars was differing beliefs amongst health providers about the risks of future disease. They examined dental practitioners' judgements on the risk of future pathology associated with pathology-free asymptomatic third molars in two studies (Kostopoulou *et al* 1998, 2000). Ten oral and maxillofacial surgeons and 18 general dental practitioners whose experience ranged from 5-28 years were presented with periapical radiographs of 36 asymptomatic, disease-free mandibular third molars with a brief

history. In the first study (Kostopoulou *et al* 1998), participants were asked to assess, using visual analogue scales, the likelihood of future pathology if the third molar was left in situ. To assess intra-observer reliability, 36 cases were duplicated and presented to the participants on a different occasion one month later. In the second study (Kostopoulou *et al* 2000), they were asked to assess the likelihood of future pathology in general and, more specifically, of root resorption, pericoronitis, periodontitis, cystic change and neoplasia. A significant variation was observed between examiners in both studies but variations between the two groups were not significant. Authors concluded that treatment decisions were not made on a rational basis (Kostopoulou *et al* 1998). Practitioners varied very considerably in their judgement of the risks of pathology associated with asymptomatic disease-free third molars and specialisation did not account for this variation (Kostopoulou *et al* 2000).

Similarly, a large degree of inter-observer variation in decision-making and treatment planning has previously been reported in other groups of dental experts (Grondahl 1979, Elderton and Nuttall 1983, Reit and Grondal 1984, Reit and Grondal 1988).

For instance, in the field of restorative dentistry, Elderton and Nuttall 1983 studied the agreement between 7 general dental practitioners and 8 dentists working within a dental hospital environment when treatment was planned for the same group of 18 young adults. The number of tooth surfaces planned for restoration by the different dentists ranged from 20 to 153. More than half the dentists agreed upon only 41% of the restorative treatment decisions made. Furthermore, it was calculated that a second dentist would only agree to fill four out of every ten surfaces that were planned for

restoration by a first dentist. This study found that the general dental practitioners tended to plan more restorative treatment than did the hospital dentists. However, there was greater agreement between the GDPs as to which tooth surfaces required treatment. This suggested that the majority of restorative treatment received by young adult patients during a single course of treatment was the result of idiosyncratic decision-making. Elderton and Nuttall concluded that a great deal of restorative treatment provided at that time was the result of "grey area" decision-making.

2.5 Making clinical decisions

Traditionally, clinical decisions have been considered an intuitive process whereby the clinician combines information about a patient and test results and available evidence to make a diagnosis or treatment plan. However, judgement analysis studies have shown that people often do not actually use the information and evidence which they consider beforehand as important when making clinical judgements (Kostopoulou *et al* 1998 *op. cit.*).

Therefore, there are other factors entering into the process of deciding the appropriate management of patients. This results in variations in decision making amongst clinicians.

Kay and Nuttall (1997) stated that variation in decision-making among clinicians could come from two main sources:

Perceptual variation (when things are perceived differently by different people) and judgemental variation (as a result of differing values about what constitutes a positive outcome).

Perceptual variation

Perceptual variation is random in nature and is an active process. It is not predictable and is less persistent than judgemental differences. Reviewing and double-checking can most likely prevent perceptual variation.

Past experience plays an important role in most individuals' behaviour: their decisions can be influenced by what happened in the past. This reinforces a person's awareness of pre-formed perception, in other words a person becomes "hyper-perceptive" of past clinical experiences. These sets of ideas are formed at a subconscious level, based on cumulative past experiences. This will in itself alter the perception of a disease process in a complex manner.

Judgemental variation

On the other hand judgemental variation is predictable and therefore can be expected to be amenable to modification. This would become possible by establishing specific criteria which would probably result in the limitation of observed variations in decision-making.

Decision-makers can be trained to use such criteria more efficiently and more reliably. The setting up of strict guidelines in different areas of dentistry would encourage dentists to make similar and more coherent judgements.

Variation in judgement among clinicians is a natural and acceptable phenomenon provided there is a rational basis for the decisions made. This could be expected to result in reliable, high quality treatment planning. This necessitates an increase in choices that will in turn increase the chance of a favourable outcome, while at the same time reducing unfavourable results (Kay and Nuttall 1997).

2.6 Chapter Conclusion

As a part of an evidence-based and patient-centred approach in clinical practice, the collection of relevant research findings and development of valid recommendations appears to constitute a fundamental step in evidence-based clinical decision-making and in the provision and delivery of highly efficient and coherent health care.

Evidence-based practice may revolutionise the way that clinicians solve their clinical problems and provide care for their patients. However, this is likely to be achieved only if relevant research and valid guidelines are appropriately incorporated into practice.

A large volume of research evidence and clinical guidelines including different publications and electronic databases such as the Cochrane library and guidelines such as those published by NICE and SIGN are available. However, the availability of these resources does not necessarily ensure their uptake by clinicians and may not lead to any change in their traditional practice. Since passive methods of dissemination of relevant information were shown to be ineffective, there is a need for more specific and active strategies to introduce the evidence into practice and ensure practice change. Existing studies seem to suggest that more intensive methods and multifaceted strategies are more likely to be effective in altering health care practice.

In the exploration for any effective strategy to incorporate evidence and bring about a change, the identification of potential barriers or facilitators to any change - and accordingly adoption of an appropriate dissemination and implementation strategy - is essential.

Generally, it is accepted that clinical decision-making based on good quality evidence can lead to more effective and efficient patient management (Richards and Lawrence 1995). Yet evidence-based health care and evidence-based guidelines have their own limitations. In the absence of reliable long-term evidence for the most part, further studies are required to test whether or how they affect the process of care and patient outcome in different settings and circumstances. Meanwhile, as decisions in clinical practice are not based upon rational thought alone, it is important that efforts be made to improve access to evidence for health professionals at their point of decision-making and to identify the potential barriers that stand in the way of behavioural change and adaptation of evidence into daily practice.

CHAPTER THREE

Methods

3 Introduction

This chapter describes the materials and methods used in the trial to evaluate the effectiveness of different strategies for dissemination and implementation of the SIGN guideline in dentistry for "Management of Unerupted and Impacted Third Molar Teeth" (SIGN 43, 2000). This pragmatic trial, based in a primary care setting, was funded by the NHS Research and Development Programme (Trial R2-64). Ethical approval from the Multi-Centre Clinical Research Ethics Committee for Scotland (MREC) as well as twelve local health authority research ethical committees was sought and an approval was subsequently granted in 1998.

3.1 Design

In this study a cluster randomised-controlled trial with a 2x2 factorial design was adopted to evaluate the effectiveness of two interventions. This trial was designed to be a pragmatic randomised controlled trial in which interventions are compared in a realistic setting. A 2x2 factorial design was adopted because this allows evaluation of the effectiveness of each intervention on its own and also provides an opportunity to study and test for interactions between the interventions if present.

The trial was conducted through dental practices across Scotland and randomisation was carried out at the practice (cluster) level, as contamination was likely if practitioners were used as the unit of randomisation. Those who worked in the same practice and

who were assigned to different experimental groups were more likely to modify or contaminate the assessment by affecting each other. They could also discuss the information they received, with resulting contamination between experimental groups (Wilson *et al* 2000). The "cluster randomisation" was advocated to minimise this contamination between intervention and the control participants (Torgerson 2001).

3.2 Sample size and power calculation

It is estimated from the literature that between 30% and 60% of extractions are inappropriate (Brickley *et al* 1996, II) and national figures for third molar extraction in dental practices suggested that 4-5 patients per practice would have extractions during each data collection period (pre- and post-intervention). This estimate was based on calculations informed by the "Scottish Dental Practice Board Mirror (SDPBM)" sited at the Dental Health Services Research Unit (DHSRU) (Pitts *et al* 1997). The SDPBM provides access to all GDP payment claims by item of service for Scotland.

For a cluster randomised trial, there is some loss of power due to the randomisation by cluster rather than individuals and this should be reflected in the sample size calculations (Kerry and Bland 1998). To ensure that the study achieves the required sample size and power for a cluster trial, an intra-practice (intra-cluster) correlation coefficient (ICC: a statistical measure of the degree of correlation within cluster) of 0.1 was incorporated into the power calculation to account for the unit of analysis being dental practices rather than individual practitioners (Donner *et al* 1981). For calculating the sample size in this study, the outcome measure was considered to be the proportion of patients having appropriate third molar extractions. There is a lack of available data

for calculating the ICCs for this condition and a number of assumptions have been made for the sample size calculation, assuming the worse scenario of an ICC of up to 0.1.

An *a priori* sample size of 60 practices, collecting information on 240 patients (minimum number of patients 240-300) was estimated to detect a 20% reduction in inappropriate extractions (based on the higher figure) from 60% to 40% assuming a 80% power and a 5% significant level (Casagrande *et al* 1978).

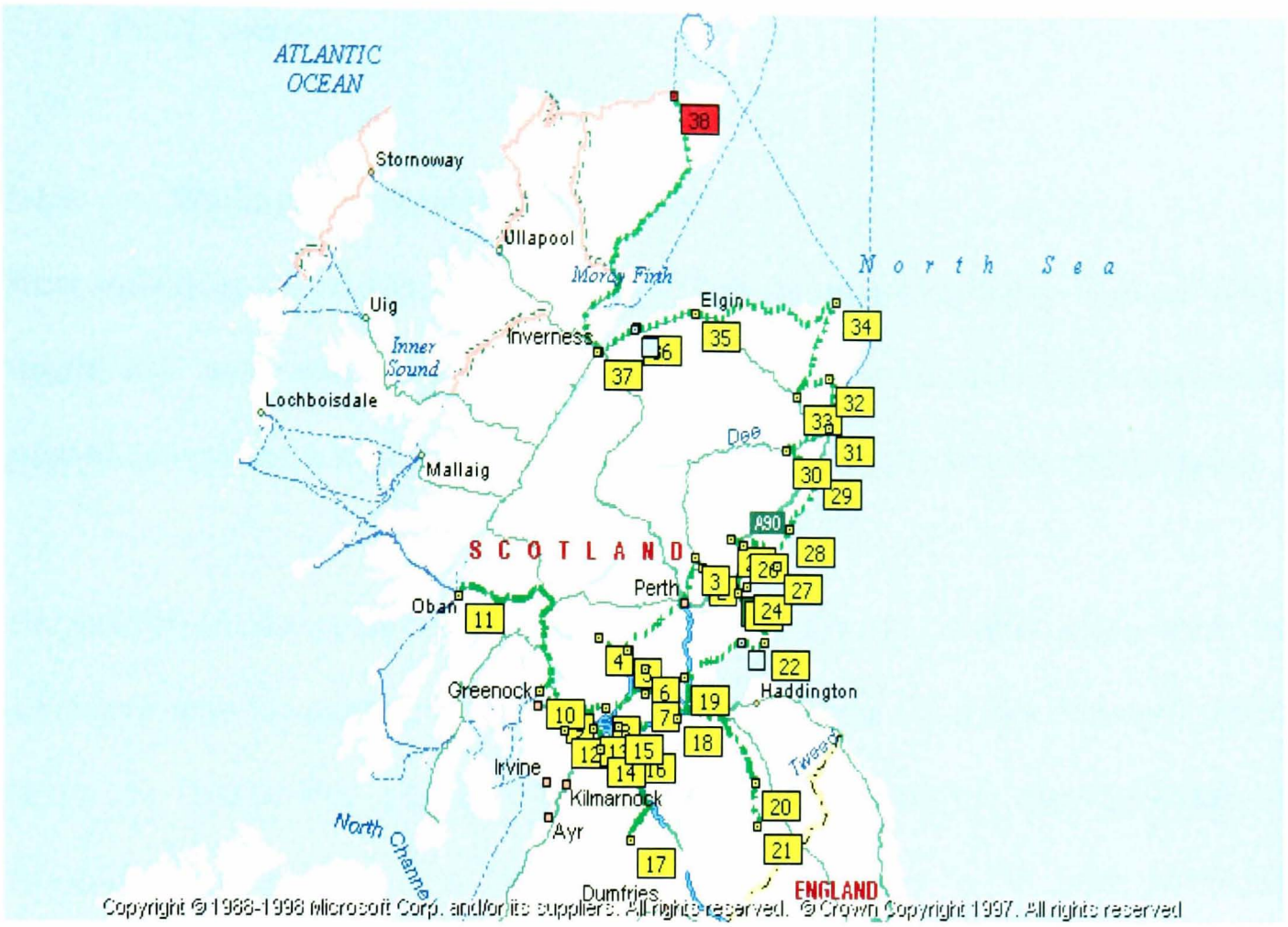
3.3 Recruitment of practices

To recruit the required sample of participants, it was necessary to identify registered dentists to whom the invitation to participate in this project could be sent. Hence, a database of all practitioners and their practice addresses from the Scottish Dental Practice Board (SDPB) in 1998 and a list of practices in the Dental Health Services Research Units Practice Based Research Network (DHSRU-PBRN) (i.e. Scottish Caries Control Study) were considered. The SDPB process the remuneration claims for dental treatment provided by general dental practitioners on the National Health Service in Scotland. Each dentist has an individual list number. The list is continually updated and corrected and can be considered as the population of general dental practitioners working for the NHS in Scotland. In addition, the Dental Health Services Research Unit also has a list of all practitioners who over the years participated in the Scottish Caries Control Study (SCCS) (Deery *et al* 1999). For this trial, dental practices across Scotland were recruited by random selection from the SDPB list, in addition to all practices in the DHSRU-PBRN list. These practices were invited to participate in the trial by mail and a recruitment of dentists was completed after four mailings. Two

mailings were made in November 1998, when 115 dental practices (include all practices in the DHSRU-PBRN list) were asked to take part, of which 33 (29%) responded positively. The third mailing went out to an additional 250 dentists in December 1998 and 10 (4%) dentists agreed to participate. The fourth mailing, in January 1999, invited another 200 dentists and positive responses from 20 (10%) of these brought the total to 63 (Figure 3.1). The recruitment letters gave details of the study and what would be involved for the participating practices (Appendix II, Letter 1 and General practitioners information sheet). The practitioners were given the option of participating individually or if other members of their practice were interested they could also become involved. Finally, if they agreed to participate, they were asked to complete a personal information sheet and consent form (Appendix II). All the forms were printed on A4 coloured papers. The colour paper was chosen to try and highlight the forms to allow easy identification by practitioners. A pre-paid reply envelope was also included.

During the period from July to August 1999, 12 practices withdrew from the project before pre-intervention data collection and randomisation of dental practices. During the period between pre-intervention and post-intervention (December 2000 to March 2001) a further four practices withdrew. As a result, the sample of participating practices was reduced to 47.

Figure 3.1 Participating practices across Scotland



3.4 Randomisation

Fifty-one practices in our study were randomised into 4 groups by a statistician independent of the research team through the computer generation of a random number sequence. The design of four intervention groups was as follows:

I Mailing of SIGN Guideline, plus postgraduate continuing education courses (PGCE) (*i.e.* non-intervention/control)

II Audit and feedback (A&F), plus mailing of SIGN Guideline and PGCE courses

III Computer aided learning with decision support (CAL-DS), plus mailing of SIGN Guideline and PGCE courses

IV Audit and feedback (A&F), plus CAL-DS, plus mailing of SIGN Guideline and PGCE courses

3.4.1 Mailing of guidelines and courses

Most guidelines are merely disseminated by mail to targeted groups. This as well as "traditional" postgraduate continuing education (PGCE) courses was used to present the guidelines to all the intervention groups as well as the non-intervention/control group.

The evidence-based guideline for the appropriate removal of third molar teeth was developed with the support of the Scottish Intercollegiate Guideline Network (SIGN) (SIGN 43, 2000). The guideline provides the broad principles which underpin the decision to remove or not to remove a third molar tooth in the form of specific recommendations. The guideline was developed following careful deliberations by the SIGN guideline development group (SIGN 43, 2000) building on existing guidelines (such as the National Institutions of Health (NIH) consensus criteria 1980 and the Royal College of Surgeons of England guideline 1997), using SIGN methodology to develop recommendations based upon the best evidence available. Their methodology involved an extensive review and appraisal of existing literature. This guideline was therefore used in this study.

Publication of the SIGN guideline for "Management of Unerupted and Impacted Third Molar Teeth" (SIGN 43, 2000) was delayed by one year but it was eventually published in March 2000. The late publication caused a delay of one year in the start of this study. The initial mailing of the guideline was the responsibility of SIGN and copies of the

guideline were distributed to all general dental practitioners in Scotland including the recruited practitioners.

Participating dentists were also invited to attend a PGCE course along the Section 63 model (i.e. postgraduate courses run by NHS Education for Scotland for GDPs) to inform them of clinical and radiological aspects of third molar teeth based on the SIGN guideline (SIGN 43, 2000). Although the courses were available to all participating practitioners, their attendance was not obligatory as this was a pragmatic trial.

Four speakers were invited to speak at each course (see timetable of the Glasgow venue as a sample in appendix III). A list of possible venues was drawn up based on the geographical position of participants. The Glasgow Dental Hospital, Aberdeen Royal Infirmary and Centre for Continuation of Education, University of Dundee were chosen, as they were most convenient and accessible for participants to attend. The courses were arranged for a half-day. Courses in Glasgow and Dundee were held in June 2000. The course in Aberdeen was cancelled due to lack of support but those who were interested in the Aberdeen course attended the Dundee course. Of 47 participant dentists who continued with the study post-intervention, 9 attended the Dundee course and 15 attended the Glasgow course. The number of participants in the courses per intervention group is presented in Table 3.1. The course attendees were asked to fill in a course evaluation and feedback form and the results of the feedback can be found in Appendix III.

Table 3.1 The number of participants in the courses per intervention group

	Groups	No of attendees in the courses (Total no of participants in each group)
Group 1	Non- intervention/control	8 (11)
Group 2	A&F	5 (12)
Group 3	CAL-DS	5 (11)
Group 4	A&F + CAL-DS	6 (13)

3.4.2 Computer aided learning with decision support (CAL-DS)

The CAL-DS was designed as a personal-based support tool, with the potential for assisting dental practitioners in deciding on the appropriate treatment of third molars. The decision support software was developed by the Department of Applied Computing, University of Dundee as part of this study. Its content was based solely on the SIGN guideline (SIGN 43, 2000) and is supported by computer-delivered advice in a multimedia format. In addition, the package contains a patient decision support system which also incorporates a multimedia information delivery sub-system to provide patients with valid information on the pros and cons of a third molar removal. However, its impact was not assessed in this trial. The use of the patient decision system was optional and was intended to save the dentist time by explaining the nature of third molar treatment and assisting with the medico-legal requirements of informed consent. Patients could explore the decision support system on their own while waiting to be seen by the dentists.

The software was piloted by a group of general dental practitioners (4 GDPs) prior to its use by participants in this study. For this part of the trial, 31 laptop computers were purchased, all of which would have been used if we had kept all of the participants on

board. However, twenty-four participants were exposed to the CAL&DS software by distributing the computers from June to September 2000. For logistical and economic reasons, some of the laptops (n=10) were distributed during the PGCE courses. Laptops were distributed to participants at the end of the courses to avoid bias and they were asked to refrain from discussing the decision supporting programme with other individuals attending the course.

The remainder of the laptops (n=13) were hand delivered to all the practices except one, which was posted due to the long distances involved. Practitioners in groups CAL-DS and CAL-DS + A&F were instructed to use the software either before, during or after a patient consultation.

3.4.3 Audit

For the audit and feedback (A&F) arm of the trial, advice was sought from the Scottish Council Dental Audit Tutor for selecting the audit facilitator and grouping.

Accordingly, the participants in this arm of trial (Groups 2 & 4) were divided into 10 groups according to distance and geographical position of their practices. The Scottish Council provided funding for the participating dentists in the audit projects.

Facilitators, whose responsibility was to organise the group, were chosen according to their previous experience and knowledge of carrying out an audit. In one group the previous experience of the participants in audit was unknown, therefore a more experienced practitioner in audit, not involved in the study, was asked to take the role of the facilitator.

Some groups were given the option of carrying out the audit in the group as allocated, or if they felt it would be difficult to meet up with the other participants because of distance, they could carry out an audit on their own with support from the project team. Or they could invite other interested practitioners in their area to collaborate in the audit. The participants of only one group decided to carry out an audit on their own which brought the total numbers of groups to 11.

Initially, all the facilitators were contacted to obtain their consent to act in this role. Then all the participants were contacted. Each group was offered two different project titles. They had the option of choosing one of these project titles or they could select any other topic relating to third molars if it was of interest to their audit group (a list of audit titles can be found in appendix IV). The design and conduct of the audit and feedback was decided within each audit group and supported by the researcher (MB) with help and advice from the Scottish Council Dental Audit Tutor.

Each group was asked to have a minimum of three meetings during their audit period. At the last meeting, the facilitator gave feedback to the group by presenting a summary of their audit results. The audit and feedback projects were conducted from June to September 2000, although 3 groups had their final meeting in November 2000.

3.5 Data abstraction

The study was a prospective cross-sectional study collecting data from patients' dental records before and after the intervention. Patients were eligible for inclusion in the study if they were 16 to 24 years old (this being the peak age for third molar problems

as third molars generally erupt between the ages of 16 and 24 years (SIGN 43. 2000)). Participating dental practitioners were asked to identify all patients in this age band who attended their dental surgery over a 4 month period prior to each period of data collection and to register their name and date of birth on the form provided (Appendix V, Record Form C).

In a separate mailing, a date for collecting data was arranged with each practice to allow the researchers to visit their practice and to carry out this task (Appendix V, Letter 2). All records were retrieved by the researchers who were blind to the intervention groups. Pre-intervention data were abstracted from 49 practices by searching dental records of all the patients registered for this project. This was done by examining paper records or electronic copies (from August to December 1999). The reason for the pre-intervention data gathering was to evaluate the current practice at participating practices prior to publication of the guideline as a baseline. This provided information about the effect of guideline dissemination on current practice as well as examining data-gathering methodology and piloting the data entry forms and methods.

Data entry forms were initially designed in a paper format. The type of data collected was determined from the SIGN guideline (SIGN 43, 2000). However, the reason for attendance, diagnosis of pathology, treatments prescribed and other clinical information were also recorded (Appendix VI, Data entry form).

The use of the electronic data entry form was planned from the start of the data collection, however, the electronic format was not ready and the paper format had to be

used for only a few (9) of the practices during the first phase of the study. The data from other practices were entered by direct entry into the computer using SPSS (Statistical Package for the Social Science) Builder Data Entry Programme 1999.

Early on in the study and before the data collection period, in May 1999, the initial clinical researcher (JD) resigned and another researcher (MB) was employed. Since she was unable to start on a permanent basis till October 1999, a temporary researcher (DG) was employed from July to September 1999 to avoid any delay in the progression of the trial. Hence, pre-intervention data were collected from 23 dental practices by the temporary researcher (DG) from August to September 1999 and data from the remaining practices (26) were collected by the permanent researcher (MB) from October to December 1999. In order to ensure a consistency in the format of data gathering, DG trained MB in data abstraction and they collected data together from two practices.

This process was repeated following the publication of the SIGN Guideline. Data were collected from 47 practices for the post-intervention stage by the permanent researcher (MB) from February to June 2001. The data collection form was not altered from pre- to post-intervention in order to have consistency in the information collected which makes the comparison between data collected from both phases of the study easier and eliminates "instrument bias" (Cook and Campbell 1979).

Following the completion of each data gathering a thank you letter incorporating a £50 gift voucher was sent to all participants to thank them for their contribution to the project.

To ensure that data involving patients' information was protected by reasonable security, safeguards against loss, unauthorised access or disclosure and other misuse, all the paper data was locked in a filing cabinet and all of the electronic data were secured with a password. The reason that patient details such as names were collected was to prevent any mistake in the data collection such as information duplication. The patients' details were deleted afterwards in order to keep the data analysis anonymous.

3.6 Statistical considerations

3.6.1 Implications of cluster randomisation for analysis

In this study, the cluster-randomised design had a number of implications for data analysis. It was assumed when practices were randomised, the patients' outcomes would differ less within the individual practice and that patients in the same practice were not independent units for the purpose of analysis (Campbell *et al* 2000).

There are two general approaches for analysing the data from a cluster randomised controlled trial (Campbell *et al* 2000). The first approach is to analyse by the unit of randomisation (cluster level analysis) i.e. practice. A summary measure such as a mean for each cluster is calculated and uses simple statistical tests for significance such as the t-test. The second approach is to analyse at a lower level than the unit of randomisation i.e. patients (patient level analysis) using multivariate models which can include the

cluster effect, the effect modifiers and relevant interactions. This form of the regression models allows for a common methodological procedure in this type of study.

When the data have been collected in two phases *i.e.* baseline and post-intervention, two methods of analysing this type of data have been suggested for whatever the unit of randomisation (i.e. cluster or patient level) is. First, the procedure is to examine the baseline imbalances. If there are no notable baseline imbalances, the post-intervention data is analysed for any differences, knowing that the groups under comparison are balanced for any confounding factors at the baseline. The second procedure examines the differences between pre- and post-intervention. Our data and analysis allows us to determine the implication of these two empirical strategies.

3.6.2 Analysis

All the analyses reflect the design of the study. A "P-value" at the 5% ($P<0.05$) was considered as statistically significant throughout the study.

3.6.2.1 Assessment of compliance with guidelines (EBP)

Adherence to the Guideline or Evidence-Based Practice (EBP) was assessed by examining each case and comparing it with the guideline. If cases were treated according to the guideline, the allocated code was 1, otherwise it was 0 (for being different). Each recommended point in the guideline was initially coded and these codes recorded for each case to indicate the reason behind the decision. The rationale was to verify the underlying principle of the chosen outcome measured by each assessor.

Data were analysed in SPSS for Windows 10 and the descriptive statistic was examined for accuracy of data entry and description of the data. Then the principle outcome measure, which was the number of patients whose treatments complied with the guideline, i.e. treatments that constituted Evidence-based Practice (EBP), was assessed. The treatments provided by participating practitioners for patients with third molar problems at pre- and post-intervention were scrutinised and evaluated independently by two clinical researchers (MB, CD) who were blind to intervention groups, in order to determine whether dentists complied with the guideline or not. Disagreements were discussed and concurrence achieved.

In addition to the important recommendations by the SIGN guideline (SIGN 43, 2000) about different aspects of third molar management, several points were considered individually by each researcher to assist with the decision-making and evaluation of the outcome measure. The deliberated points are set out as an example, in Table 3.2.

Table 3.2 **The points which were considered individually by each researcher to assist with the decision-making and evaluation of the outcome measure.**

The age of the patient at the time of the recorded problem with the third molar tooth/teeth

⬇

Initial person who noticed the problem first *i.e.* dentist/ patient/ general medical practitioner

⬇

Assessment of recurrence status of the third molar tooth/teeth pathology

⬇

Assessment of whether the patient previously attended the practice with the same problem or not

⬇

Assessment of recorded clinical signs or symptoms such as pain, facial swelling or trismus

⬇

Whether any antibiotics had been prescribed or not

⬇

Assessment of the orientation and the shape of the third molar tooth/teeth in the mandible (the possible successful eruption of the tooth into a normal position was judged with the help of recorded data from radiographs taken by GDPs).

⬇

Assessment of any history of Tempero-mandibular joint (TMJ) dysfunction syndrome

At the completion of the pre-intervention data analysis, these points were discussed by the two researchers and used as inclusion criteria for the post-intervention data analysis.

3.6.2.2 Correlation between pre and post intervention

Analysis of a correlation between pre- and post-intervention cluster level compliance (EBP) rates was carried out which demonstrated a weak correlation (Pearson's Product Moment Correlation = -0.125, n = 43, t = 0.81, P=0.423) (Table 3.3, Graph 3.1). This was also true when pre- and post-intervention cluster level compliance (EBP) rates were broken down into the intervention groups (Tables 3.4, Graphs 3.2).

In addition, visual inspection of the weighted pre-intervention means did not highlight any notable baseline imbalance (see Chapter 4, Table 4.1.10). Furthermore, the multilevel model (see Chapter 3, Section 3.6.2.3b) originally used all data with a variable to indicate whether a case was pre- or post-intervention and the result was non-significant. Therefore, all the main analyses were performed on post intervention compliance rates (EBP) only (see Section 3.6.2.3).

Table 3.3 Correlations between pre and post-intervention cluster level Compliance (EBP)

		Pre- intervention EBP	Post-intervention EBP
Pre	**Pearson Correlation**	1.000	-0.125
	P-Value		0.423
	N	43	43
Post	**Pearson Correlation**	0.125	1.000
	P-Value	0.423	
	N	43	43

Graph 3.1 **Scatter plot of pre against post intervention cluster level compliance (EBP) rates**

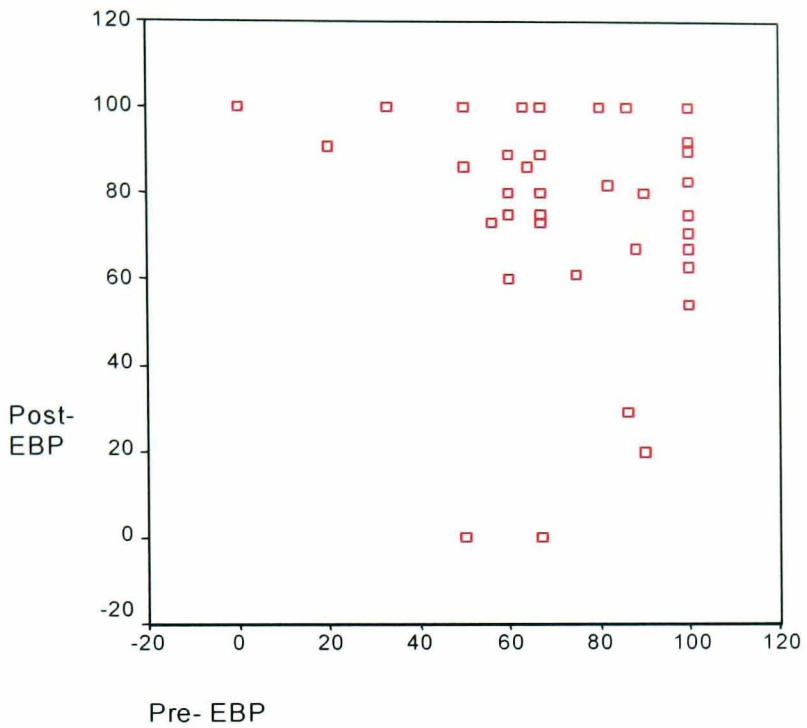
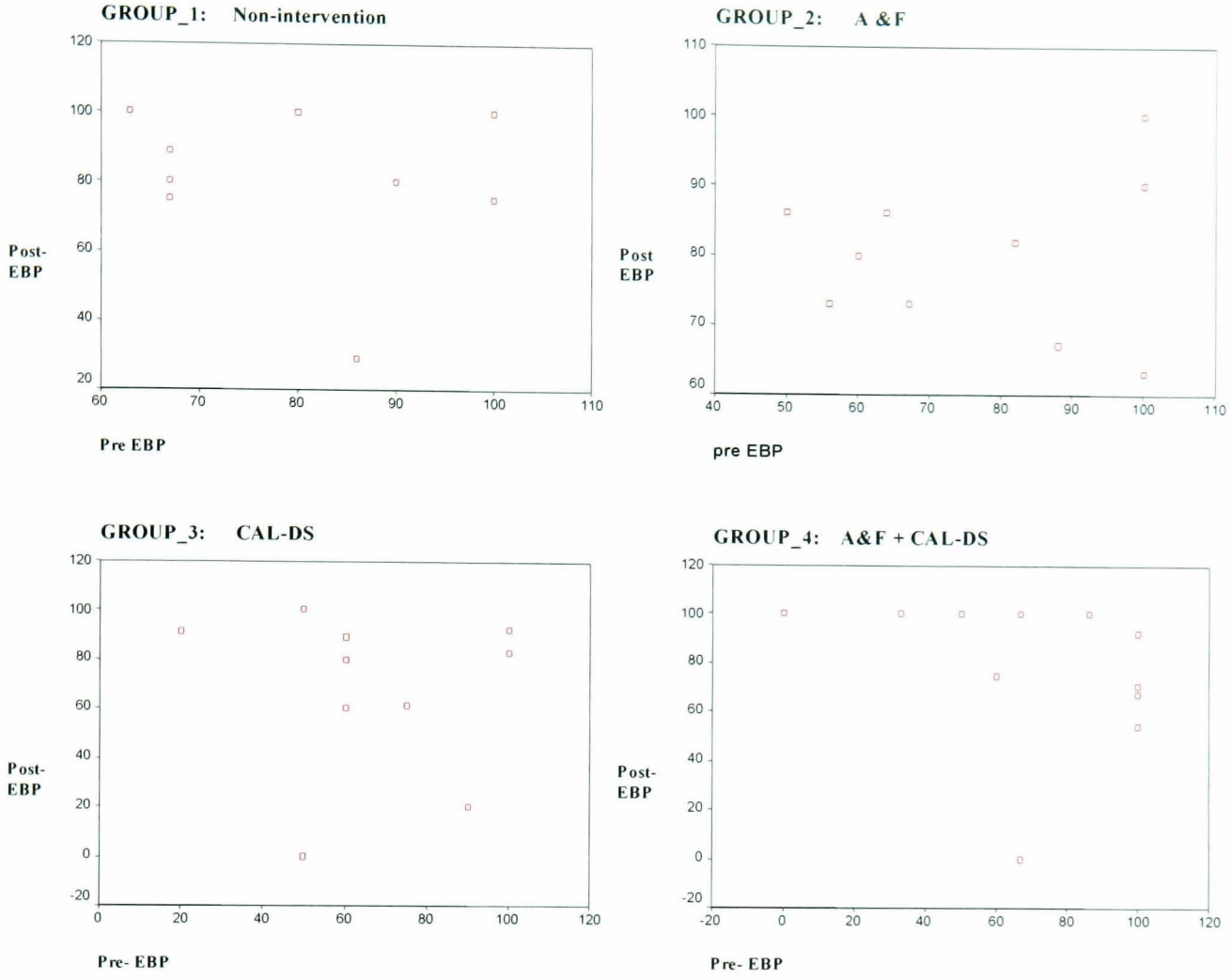


Table 3.4 **Correlations between pre and post-intervention cluster level compliance (EBP) for different intervention groups**

Intervention Groups			Pre- intervention EBP	Post-intervention EBP
Non-intervention	Pre	Pearson Correlation	1.000	-0.153
		P-Value		0.695
		N	10	9
	Post	Pearson Correlation	-0.153	1.000
		P-Value	0.695	
		N	9	11
A&F	Pre	Pearson Correlation	1.000	0.071
		P-Value		0.825
		N	13	12
	Post	Pearson Correlation	0.071	1.000
		P-Value	0.825	
		N	12	12
CAL-DS	Pre	Pearson Correlation	1.000	-0.069
		P-Value		0.850
		N	11	10
	Post	Pearson Correlation	-0.0.69	1.000
		P-Value	0.850	
		N	10	11
A&F + CAL-DS	Pre	Pearson Correlation	1.000	-0.333
		P-Value		0.290
		N	13	12
	Post	Pearson Correlation	-0.333	1.000
		P-Value	0.290	
		N	12	12

Graph 3.2 **Scatter plot of pre- against post-intervention cluster level compliance (EBP) rates for different intervention groups**



3.6.2.3 Analysis of the main effects of the interventions based on the post-intervention data

The primary analysis of the data examined the main effects of the interventions (first: CAL-DS versus no CAL-DS and second: A&F versus no A&F). For the purpose of analysis of the main effect of the interventions under investigation, the factorial design implied that the effect of CAL-DS was tested by combining Group 1 with Group 2 and comparing this to the combination of Group 3 with Group 4. Similarly, the effect of A&F was tested by combining Group 1 with Group 3 and comparing this to the combination of Group 2 with Group 4 (Table 3.5).

Table 3.5 Intervention Groups (Factorial Design)

Group	Allocated intervention
Group 1	Non-intervention/control
Group 2	A&F
Group 3	CAL-DS
Group 4	CAL-DS + A&F

3.6.2.3a Cluster level analysis

The level of compliance with guidelines (EBP) between intervention groups was analysed by means of a cluster level analysis using a weighted t-test. Using a standard two-sample t-test weighs each practice estimate as the same. Therefore the practice compliance estimates were weighted by the number of patients seen in that practice.

In addition, in order to control for heterogeneity in patients (case-mix) and to address the hierarchical nature of the data (patients nested within dental practices), hierarchical (multilevel) analysis was also adopted.

3.6.2.3b Hierarchical analysis

A multilevel logistic regression was adopted as the technique to analyse the datasets using the multilevel modelling software MLwiN for windows: release 6.0. This technique allows for the effect of confounding factors such as heterogeneity in patients (case-mix) to be accounted for in the analysis. For example three conditions (pericoronitis, caries and pulpal pathology) were considered to be potential effect modifiers (confounding variable), as they are exogenous variables and are not within the control of the dentist. Every patient who presented with pulpal pathology (100%) was

treated in accordance with the guideline. For patients presenting with caries, only one case in the CAL-DS group was not treated in accordance with the guideline. Therefore, there was no advantage gained in investigating these two conditions further. The only disagreement with the guidelines came when patients presented with pericoronitis. Therefore pericoronitis was considered as an effect modifier (Chapter 4, see Section 4.2.1).

Primarily, all the possible variables and interactions between time, intervention and patient characteristics or case-mix were modelled. However, the difficulty was with the low number of cases. Consequently, the cells in the regression model became very small and many of them were empty which can result in unreliable standard errors. Therefore, when the model was considered it was necessary to discard all cases with a low number.

In order to determine the main effect of the intervention groups (A&F and CAL-DS) on EBP, a regression equation was estimated. This assumed a binomial error structure. The subscripts i and j denote the patient and practice. The expected proportion π_{ij} denotes the probability that the i^{th} patient's treatment in the j^{th} practice complied with guidelines (EBP). The probability π_{ij} is modelled using a logit function that consists of an intercept term β_{1j} and explanatory variables. For a single explanatory, β_2 (for example, the effect of Audit), the estimated equation is

$\text{logit}(\pi_{ij}) = \beta_{1j} + \beta_2 x_{2ij}$ where $\beta_{1j} = \beta_1 + u_j$. The intercept was random at the level of the dental practice. The main effects of CAL-DS and A&F were modelled as fixed

effects. The advantage of this method is that individual patient characteristics can be modelled simultaneously (e.g. if a patient has pericoronitis).

The interaction effects were modelled as.

$$\text{logit}(\pi_{ij}) = \beta_{1j} + \beta_2 x_{2ij} + \beta_3 x_{3ij} + \beta_4 x_{4ij}$$

Where, x_{2ij} equalled 1 if the patient was in an intervention group that received A&F (i.e. group 2 or 4). x_{3ij} equalled 1 if the patient was in a group that received CAL-DS, (i.e. group 3 or 4). These two terms model the main effects. x_{4ij} equalled 1 if the patient received both CAL-DS and A&F, that is to say, was in group 4. This term models the interaction effect.

3.6.2.4 Sub-analysis : Change of EBP within the groups post-intervention

To examine the change in EBP within the groups from pre- to post-intervention, a sub-analysis regression model was adopted using the Stata statistical software for windows: release 6.0. In order to determine the net effect of the intervention groups on EBP, a regression equation as estimated (Eq 3.1).

Eq 3.1

$$y_{ij} = \mu + \sum_{k=2}^{4} G_{ijk}\alpha_k + \sum_{m=1}^{3} X_{ijm}\beta_m + T_{ij}\gamma + \sum_{k=2}^{4} (G_{ijk}T_{ij})\alpha'_k + \sum_{m=1}^{3} (X_{ijm}T_{ij})\beta'_m + \nu_j + \varepsilon_{ij}$$

Explanatory variables incorporated into this model are: The first group of variables include design factors such as intervention groups. The second group consists of patients' characteristics or "confounding/case-mix variables" (*i.e.* referred, caries, and

pericoronitis) that may modify the effect of the interventions. The referred variable was used as a proxy for the difficulty of the case and was not assessed previously in the main analysis. The third group of variables consisted of a time variable denoting pre- or post-intervention. The fourth group of variables related to the interaction terms between time variable, confounding variables and intervention variables.

The following notation is used in the equation (Eq 3.1). The subscripts i, j, k and m denote the patient, practice, intervention group and case-mix control (patients' characteristics) variables, respectively. y_{ij} represent the outcome of the i^{th} patient in the j^{th} practice in which it is equal to one if the treatment agrees with the guideline (EBP) or zero if it is otherwise. The variables G_{ijk}, X_{ijm} and T_{ij} are explanatory variables taking the value 1 or 0, if the i^{th} patient with m characteristic, in j^{th} cluster/ practice, belonging to the k^{th} intervention group at the T^{th} time point (T=0 for baseline, T=1 for post-intervention). μ is the intercept or constant, β is the regression coefficient describing the relationship between the outcome (y_{ij}) and the covariate (explanatory variable). ν_j is the random effect (level 2) for the j^{th} practice/cluster and ε_{ij} (level 1) is the residual for the i^{th} individual patient in the j^{th} practice/cluster. The random effect ν_j represent the amount by which the intercept for the j^{th} cluster differs from the overall mean value μ. The dependence between observations within the same cluster is modelled explicitly via the random effect ν_j. It is the presence of the two variance terms ν_j and ε_{ij} that defines the model as a multilevel or random effect model.

There are two important issues raised by the dependent variable in this study: first, the outcome variable (yij) is binary (1 if EBP or 0 otherwise); second, our data is based on multiple observations per practice, hence a multilevel model.

All statistical tests of interest are linear restrictions of equation 3.1.

3.6.2.5 Descriptive analysis: change in the proportion of "patient" variables after interventions

Individual variables were examined for the possible descriptive patient outcomes, using a Chi-square statistical analysis to determine the changes in the proportion of "patient" variables before and after the interventions. Log-linear analysis such as HILOGLINEAR in SPSS for Windows 10 was used to summarise and highlight the associations in complex cross tabulation tables.

The treatments recommended by the SIGN guideline (SIGN 43, 2000) can be categorised into four broad categories: refer, extract, observe, and restore. All the conservative treatments such as mouthwash, local irrigation, and prescribed antibiotics were grouped together as observation. In analysis, treatments were categorised accordingly (Chapter 4, Section 4.4.1.2). Treatments were further divided into two broader groups *i.e.* "reactive" and "proactive" (Section 4.4.1.2, 4.4.1). "Proactive" treatments were defined as treatments that involved removing a tooth or referring a patient to a specialist clinic. "Reactive" treatments were defined as conservative treatments such as restoration, observation including mouthwash, local irrigation and prescribed antibiotics.

Finally data for removal of the third molar teeth by Scottish dentists were examined to assess the change in rate of removal of these teeth in a designated population of practitioners using data from NHS Scotland's Management Information and Dental Accounting System (MIDAS).

Given the pragmatic nature of the study, the analysis was performed in terms of an intention to treat (ITT). All patients managed by the study dental practices were analysed as study patients whether or not the practitioners fully utilised the guideline implementation strategies (Newell 1992, Prescott 1999).

3.7 Protocol deviations

Participating practitioners were asked at the beginning of the study to record the names of their patients aged 16 to 24 (per protocol) for data collection. However this criterion was not fully met by the participating dentists as they registered patients aged between 14 and 25. Since analysis was by "intention to treat" these cases were not excluded from the study, as there was no reason to think that patients would be treated differently. The frequencies of the different age range before and after intervention presented in Table 3.6.

Table 3.6 Number of patients in different age range before-and after Intervention

Age range	Pre-intervention N=244 n(%)	Post-intervention N=426 n(%)	Total N=670 n(%)
14 - <16	2 (1)	0	2 (1)
16 - 24 (per protocol)	206 (84)	355 (83)	561 (84)
>24 - 25	36 (15)	71 (17)	107 (15)

N=total number of patients, n=total number of patients in each category

CHAPTER FOUR

Results

4 Introduction

This chapter presents the results of the study in several sections: Section 1 (4.1) outlines the descriptive results and demonstrates the main outcome measurement of the study as assessed by two independent researchers. In the subsequent section (4.2) the main study question is considered and the post-intervention data are examined by cluster and hierarchal analysis. In Section 3 (4.3) the pre- and post-intervention data are examined as sub-analysis in order to evaluate the relative changes in EBP within the groups following the interventions at "patient" level. In Section 4 (4.4) the data are analysed for possible descriptive patient outcomes. Any changes in the relative proportion of each patient's variables following the interventions are demonstrated as the descriptive patients' characteristics analysis. Data are also analysed by taking into account the SIGN guideline (SIGN 43, 2000) recommendations about recording medical history and radiographic examination. In the penultimate section (4.5) the rate of removal of third molar teeth by Scottish dentists is examined. This is followed by a summary of the findings in the last section (4.6).

4.1 Section 1:

4.1.1 Descriptive results

Sixty-three (11%) of the 565 invited dental practices volunteered to participate in the trial and they were asked to complete a personal information sheet and consent form. Twelve of these withdrew before pre-intervention data were collected. The 51 remaining dental practices were randomised into four groups:

Group 1: Non- intervention/Control

Group 2: Audit and feedback (A&F)

Group 3: Computer aided learning with decision support (CAL-DS)

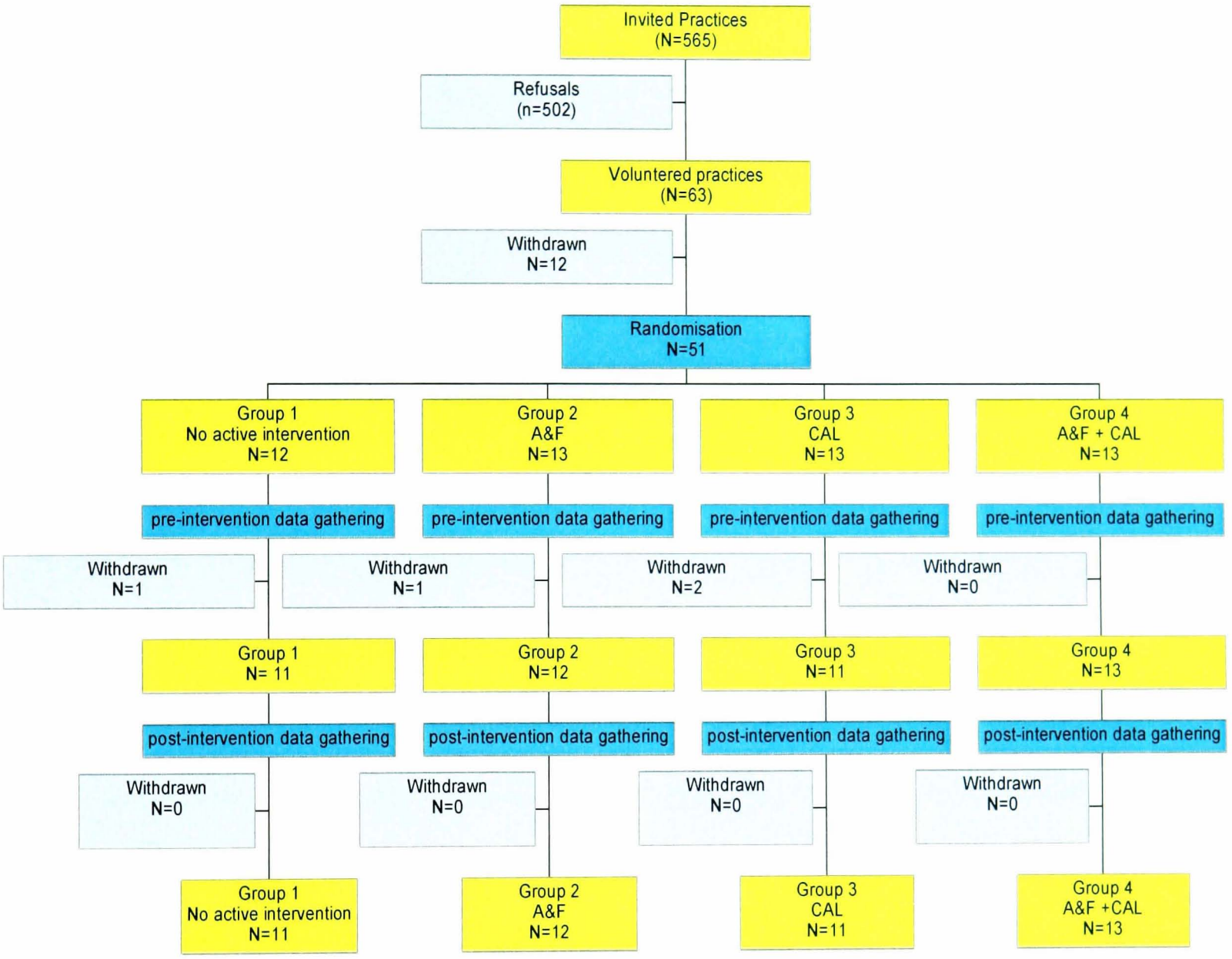
Group 4: A&F + CAL-DS.

Between the two data collection periods an additional four (8%) practices withdrew from the trial leaving 47 practices to complete the trial (Fig 4.1). In one practice the original dentist who initially agreed to participate in the trial left the practice following the pre-intervention data collection and a substitute dentist agreed to continue with the project. Since this was a pragmatic trial and analysis was by "intention to treat" (ITT), the data collected from this particular practice was treated as a no change (*i.e.* assumed continuity) from pre- to post-intervention.

Figure 4.1

Trial Profile

A&F = Audit and Feedback, CAL= Computer assisted learning



Fifty one practices were randomised, only 49 practices collected pre-intervention data and, of these, 47 had relevant data, *i.e.* third molar cases.

Post-intervention, 47 practices collected data of which 46 had relevant data, i.e. third molar cases.

The number of practices (clusters) per group is shown in Table 4.1.1

Table 4.1.1 The number of practices (clusters) with relevant data per intervention group

Groups		Pre-intervention No of Clusters	Post-intervention No of Clusters
Group 1	Non- intervention/Control	10	11
Group 2	A&F	13	12
Group 3	CAL-DS	11	11
Group 4	A&F + CAL-DS	13	12

4.1.1.1 Participating dentists profile

The profiles of the participating dentists are reported in Table 4.1.2. The mean age of principal dentists from 51 practices was 42 years (SD 7.8 years, Min=26 years, Max=58 years). This included 41 males and 10 females. The year of qualification of these dentists ranged from 1965 to 1997 (Median 1981), (Table 4.1.2). Twenty-six percent (n=13) of the participants had been qualified less than 16 years, 50% (n=26) had been qualified between 16 and 24 years, and 24% (n=12) had been qualified over 25 years.

Table 4.1.2 Dentists' characteristics overall and per intervention group

Groups		Gender Male	Gender Female	Mean age	Year of qualification
Overall		41	10	42 (SD 7.8)	1965-1997 (Median1981)
Group 1	Non- intervention /control	8	4	44	1971-1989
Group 2	A&F	12	1	38	1966-1987
Group 3	CAL-DS	9	4	44	1968-1997
Group 4	A&F + CAL-DS	12	1	41	1965-1993

There were no significant differences between the dentists who withdrew from the study and those who continued in terms of their age (t=-1.34, P=0.2); gender (χ^2=0.15, df=1, P=0.7); postgraduate qualifications (χ^2=2.24, df=1, P=0.1) and their intervention group (χ^2=4.24, df=1, P=0.2).

4.1.1.2 Patients profile

Data were collected for 3342 (1885 males, 1457 females) patients with a mean age of 21.7 years (SD 2.2, range 14 - 25.5 yrs) pre-intervention, compared with 1935 (880 males, 1055 females) patients with a mean age of 21.8 years (SD 2.1, range 16.6-25.8 yrs) post-intervention (Table 4.1.3).

Table 4.1.3 Patients' Characteristics

Pre-intervention			Post-intervention		
Male	**Female**	**Mean Age (SD, Range)**	**Male**	**Female**	**Mean Age (SD, Range)**
1885	1457	21.7 (2.2, 13.4-25.5)	880	1055	21.8 (2.1, 16.6-25.8)

There was a variation in the percentage of patients (15%) who presented with third molar problems from the pre- to post-intervention periods. The proportion of patients with third molar problems pre-intervention was 7% (n=244) compared with 22% (n=426) post-intervention. At the same time, a reduction in the overall number of patients identified for the study by the practices was observed from pre- to post-intervention (3342 pre- compared with 1935 post-intervention) (Table 4.1.4).

Table 4.1.4 The gender distribution of patients with and without third molar problems before and after intervention

	Pre- intervention (Overall no of patients) N=3342				Post- intervention (Overall no of patients) N=1934			
Patients	**Male n(%)**	**Female n(%)**	**Total n(%)**	***P value**	**Male n(%)**	**Female n(%)**	**Total n(%)**	***P value**
with third molar problem	98 (40)	146 (60)	244 (100)	0.0001	161 (38)	265 (62)	426 (100)	0.0001
without third molar problem	1787 (58)	1311 (42)	3098 (100)		719 (48)	789 (52)	1508 (100)	

*Pearson Chi-square

4.1.1.3 Gender distribution of patients with different types of third molar pathology

The data were examined to assess whether there was a variation in the proportion of patients with different pathologies in their third molar teeth, *i.e.* caries and pericoronitis, according to gender.

At pre-intervention, 60% of females presented with third molar problems compared with 40% of males. A similar distribution was found in post-intervention, *i.e.* 62% versus 38% (P=0.0001) (Table 4.1.4).

The Log linear tests of association confirmed the above findings and revealed that gender ratio and proportions of patients with third molar problems changed from pre- to post-intervention (*i.e.* represented as "time" in the table 4.1.5) (P=0.0001). Overall, females have third molar problems more often (P=0.0001). The effects of gender and time on the presence or absence of the third molar problems are independent and there was no significant association observed between gender and the third molar problem from pre- to post-intervention (P=0.09) (Table 4.1.5).

Table 4.1.5 Tests of association (Hierarchical Log linear) of third molar problem with Time (*i.e.* pre- to post-intervention) and gender

	DF	P value
Time* gender	1	0.0001
Time* third molar problem	1	0.0001
Gender* third molar problem	1	0.0001
Time* gender* third molar problem	1	0.09

Time = pre- and post-intervention comparison

More females presented with pericoronitis (62%) as compared to the males (49%) and this was statistically significant (P=0.001). On the other hand, a higher number of males presented with caries (39%) as compared to the females (26%). This was also statistically significant (P=0.001) (Table 4.1.6).

Table 4.1.6 Comparison of gender in each pathology category

Number (percentages) of each pathology				
Gender		**Male N=259 n (%)**	**Female N=411 n (%)**	***P-value df=1**
Pericoronitis	Yes	126 (49)	254 (62)	0.001
	No	133 (51)	157 (38)	
Caries	Yes	100 (39)	108 (26)	0.001
	No	159 (61)	303 (74)	

*Pearson Chi-square, N=overall number of males/females in each pathology category

It was investigated whether the above findings differed from pre- to post-intervention. The comparison of pre- and post-intervention data revealed that there were no significant changes in the proportion of females with either of these pathologies, *i.e.* pericoronitis or caries (Table 4.1.7). This finding was similar in males presenting with caries. However, data suggested a significant reduction in the proportion of males with pericoronitis presenting pre- (57%) to post-intervention (44%) (P=0.04) (Table 4.1.7).

Table 4.1.7 Comparison of pathologies for each sex before and after intervention

Gender	Pathologies	Pre-intervention n/N (%)	Post-intervention n/N (%)	Total n/N (%)	*P value df=1
Male	Pericoronitis	56/98 (57)	70/161 (44)	126/259 (49)	0.04
	Caries	37/98 (38)	63/161 (39)	100/259 (39)	0.9
Female	Pericoronitis	96/146 (66)	158/265 (60)	254/411 (62)	0.3
	Caries	35/146 (24)	73/265 (28)	108/411 (26)	0.5

N= Total number of males/ females, n= Number of males or females with each pathology
*Pearson Chi-square

4.1.2 Compliance with the guideline (EBP) as assessed by two researchers

A comparison of all the cases with the guideline (SIGN 43, 2000) and an assessment of their concordance (EBP) as assessed by the 1st (and 2nd) evaluator found that 74% (68%) of cases followed EBP before the intervention as compared with 78% (75%) post-intervention. Following a full discussion of disagreements between two assessors, 74% were judged to follow the EBP before the implementation of any intervention. This value increased to 78% following the intervention phase. This rise in compliance from 74% to 78% between two stages of the study, however, was not statistically significant (P= 0.25) (Table 4.1.8).

Table 4.1.8 Assessment of guidelines compliance/EBP before and after Intervention

Evaluators	Pre- intervention n (%) N=244	Post-intervention n (%) N=426
1st researcher	180 (74)	332 (78)
2nd researcher	167 (68)	321 (75)
Final agreement	181 (74)	334 (78)

It was found that the aggregate compliance with the guideline (SIGN 43, 2000) (EBP) was high for both the pre- and post-intervention phases (Table 4.1.8).

Compliance with the guideline (EBP) was compared for different intervention groups and these values are outlined in Table 4.1.9.

Table 4.1.9 Degree of compliance with the guidelines (EBP) for each intervention group before and after the intervention

Groups	n=agreed cases with the guidelines/N=total no of cases per group (percentages)	
	Pre-intervention	Post-intervention
	Total no of patients=244 n/N (%)	Total no of patients =426 n/N (%)
Non-intervention/control	50/65 (76)	76/94 (81)
A&F	55/72 (76)	129/165 (78)
CAL-DS	31/47 (66)	66/90 (73)
CAL-DS + A&F	45/60 (75)	63/77 (82)

The estimates from dental practices with larger numbers of patients can be more precise. Therefore the raw data, i.e. the practice compliance (EBP) estimates, were weighted by the number of patients seen in that practice. Thus, practices with fewer patients have a lesser effect on the estimate of EBP than practices with larger numbers (Bland and Kerry 1998).

An assessment of the pre-intervention compliance rates (EBP) for all groups demonstrated that there were no apparent imbalances between the groups mean percentage compliance with the guidelines (EBP) at baseline (Table 4.1.10).

Therefore, all analysis was performed on post-intervention data to assess the main effect of the interventions (i.e. Audit and CAL-DS).

Table 4.1.10 Weighted mean percentage compliance (EBP) for each intervention group before and after the intervention

	Pre-intervention		Post-intervention	
Groups	**% (SD)**	**(95% CI)**	**% (SD)**	**(95% CI)**
Non-intervention/control	77 (12)	(70 - 85)	81(18)	(70 - 92)
A&F	77 (18)	(66 - 86)	78 (10)	(73 - 84)
CAL-DS	70 (24)	(56 - 84)	73 (25)	(59 - 88)
A&F + CAL-DS	75 (24)	(62 - 88)	82 (23)	(79 - 95)

While EBP increased in all the groups including the non-intervention/control group following dissemination of guidelines and interventions, the increase does not seem to be large (see Table 4.1.10, comparing the 95% CI). Hence, the relative changes in EBP within each group (i.e. Non-intervention/control, A&F, CAL-DS, A&F + CAL-DS) were examined to assess how large this effect is.

4.2 Section 2

4.2.1 Main analyses: Analysis of main effects of interventions based on post-intervention EBP

This section addresses our main research question: what is the effectiveness of the interventions (A&F and CAL-DS) under investigation?

To address this question, the post-intervention data were analysed firstly by cluster level analysis and then by multilevel analysis because of the hierarchical nature of the data (i.e. patients nested within dental practices).

Cluster level analysis results using a weighted t-test for the comparison of weighted mean percentage compliance with the guidelines for either A&F or CAL-DS with non-intervention/control group, demonstrated that neither the audit (P=0.62) nor the CAL-DS (P=0.76) were effective in changing the adherence of dentists to the guidelines for the management of third molar teeth (Table 4.2.1).

Table 4.2.1 Percentage compliance with guidelines (EBP) for A&F versus no A&F group and CAL-DS versus no CAL-DS group

Cluster level analysis			
Intervention Groups	**Post-intervention % (SD)**	**Post difference (95% CI)**	**P-value**
No A&F	77.2 (21.4)	2.2 (-0.82 - 12.6)	0.62
A&F	79.4 (15.1)		
No CAL-DS	79.3 (13.2)	-2.0 (-13.3 - 9.3)	0.76
CAL-DS	77.3 (23.7)		

The multilevel analysis found the odds ratio of EBP for dentists who experienced A&F versus those who did not was 1.28 (95% CI: 0.62 to 2.63). This comparison was not statistically significant (P=0.51), (Table 4.2.2).

For dentists who used CAL-DS versus those who did not use CAL-DS, the odds ratio was 0.84 (95% CI: 0.88 to 1.74). This comparison was thought not statistically significant (P=0.65), (Table 4.2.2).

Table 4.2.2 Estimated probability of compliance with guideline (EBP) for A&F versus no A&F group and CAL-DS versus no CAL-DS group

Hierarchical analysis			
Intervention Groups	**Post-intervention probability**	**Odds ratio (95% CI)**	**P-value**
No A&F	0.79	1.28 (0.62 - 2.63)	0.51
A&F	0.82		
No CAL-DS	0.82	0.84 (0.88 - 1.74)	0.65
CAL-DS	0.79		

4.3 Section 3

4.3.1 Sub-analysis I: pre-intervention versus post-intervention

This section examines the relative changes in EBP within the intervention groups from pre- to post-intervention and addresses the relative effectiveness of the interventions in different groups including or excluding the controls for the "effect modifiers" or "case-mix" before and after any intervention.

As has already been shown in the method section (Chapter 3, Section 3.6.2.4), in order to determine the net effect of the intervention groups on EBP, a regression equation (Eq 3.1) was estimated and the regression estimates were evaluated with the control for case-mix (CM) or without (NCM).

To provide estimates for the probability of compliance with guidelines (EBP) without controls for the case-mix (NCM), the estimates from the robust Probit regression were utilised. These probabilities are reported in Table 4.3.1 and Graph 4.3.1.

Table 4.3.1 Estimated probability of EBP (NCM) for each intervention group

	Pre-intervention		Post-intervention	
No case-mix (NCM)	**Prob (SE)**	**(95% CI)**	**Prob (SE)**	**(95% CI)**
Non-intervention/control	0.77 (0.05)	[0.67 - 0.87]	0.80 (0.04)	[0.72 - 0.89]
A&F	0.76 (0.04)	[0.66 - 0.85]	0.78 (0.08)	[0.62 - 0.92]
CAL-DS	0.66 (0.07)	[0.51 - 0.80]	0.73 (0.10)	[0.47 - 0.89]
A&F + CAL-DS	0.75 (0.06)	[0.63 - 0.86]	0.82 (0.08)	[0.60 - 0.92]

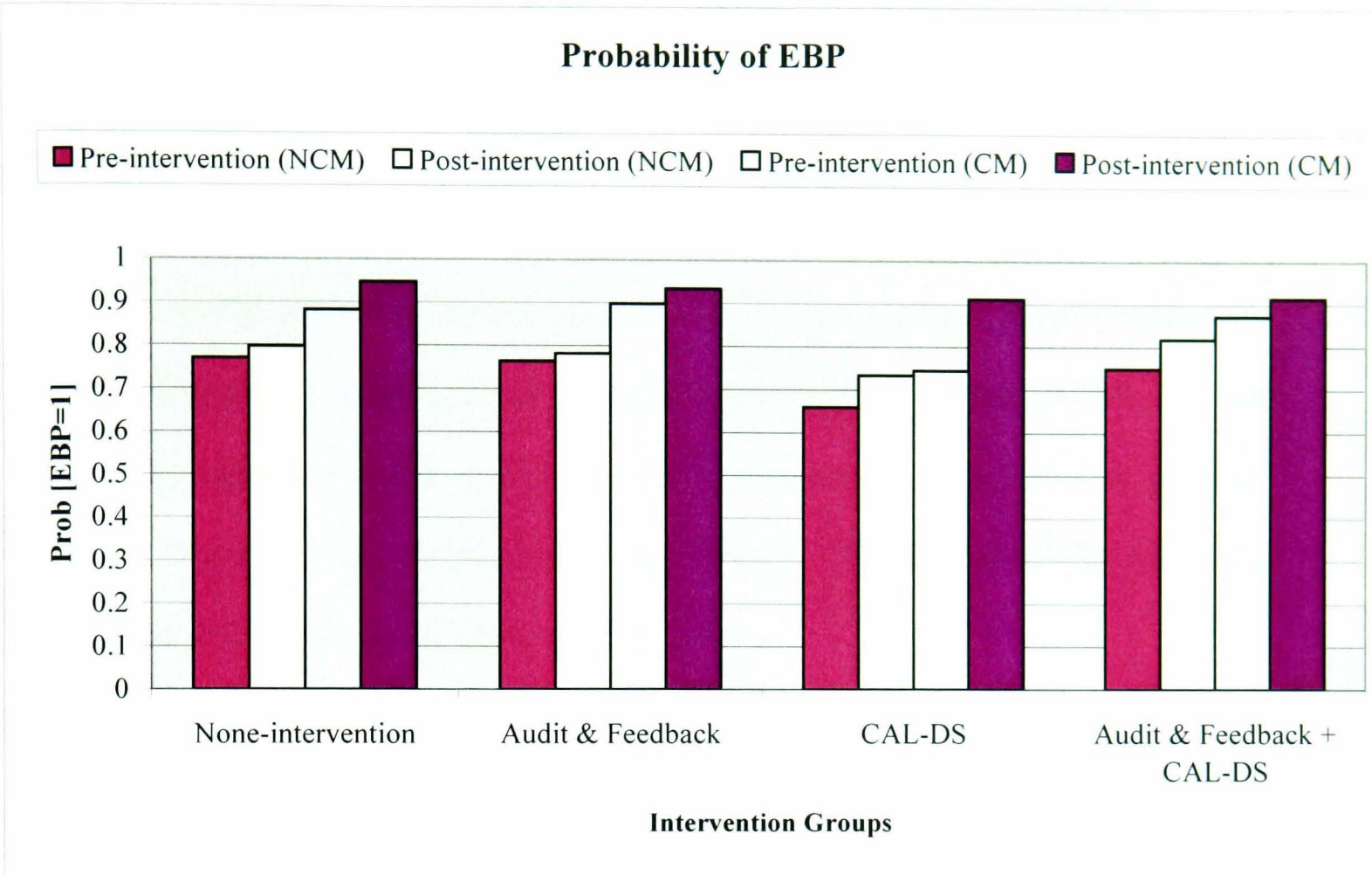
For estimates of probability of EBP including the controls for the case-mix (CM). the estimates from the robust Probit regression were employed. These probabilities are reported in Table 4.3.2 and Graph 4.3.1.

Table 4.3.2 Estimated probability of EBP (CM) for each intervention group

	Pre- intervention		Post- intervention	
Case-mix (CM)	**Prob (SE)**	**(95% CI)**	**Prob (SE)**	**(95% CI)**
Non-intervention/control	0.88 (0.06)	[0.79 - 0.99]	0.95 (0.02)	[0.90 - 0.99]
A&F	0.90 (0.05)	[0.82 - 0.99]	0.93 (0.02)	[0.89 - 0.99]
CAL-DS	0.75 (0.14)	[0.55 - 0.99]	0.91 (0.03)	[0.84 - 0.99]
A&F + CAL-DS	0.87 (0.07)	[0.78 - 0.99]	0.91 (0.04)	[0.85 - 0.99]

The estimated probability of EBP with and without controls for the case-mix (Tables 4.3.1 & 4.3.2) for different groups is shown by a graph (Graph 4.3.1).

Graph 4.3.1



In graph 4.3.1, NCM and CM denote the exclusion of control for case-mix and inclusion of control for case-mix

Graph 4.3.1 reveals a number of important points.

- First, the comparison of intervention groups at pre-intervention, when controls for the case-mix were not included, demonstrated that none of the groups differed significantly at pre-intervention (see Graph 4.3.1, Table 4.3.1, 95% CI for different groups at baseline). In contrast, there were differences between the groups when controls were included for the case-mix. Specifically, this showed a difference between the CAL-DS group and the others (Table 4.3.3). The CAL-DS was significantly different from the A&F group (P= 0.04).

Table 4.3.3 Pre-intervention comparison of estimated probability of EBP for CAL-DS versus other intervention groups including case-mix

Groups comparison	P value
CAL-DS V Non-intervention	0.10
CAL-DS V A&F	0.04
CAL-DS V A&F + CAL-DS	0.14

- Secondly, comparing the pre- and post-intervention, changes in compliance with the guideline (EBP) within the group without controls for the case-mix (NCM) were not statistically significant in any group (see Graph 4.3.1, Table 4.3.1). However, after the inclusion of controls for case-mix (CM), the change in EBP was statistically significant in the CAL-DS group only (P=0.04) (Table 4.3.4).

Table 4.3.4 Pre- versus post-intervention comparison of estimated probability of EBP for each intervention groups including case-mix

Groups comparison	P value
Non-intervention	0.1
A&F	0.4
CAL-DS	0.04
A&F + CAL-DS	0.4

- Thirdly, the changes in EBP within the groups from pre- to post-intervention did not differ between the groups when controls for case–mix were included. This has been shown for the CAL-DS group versus the other intervention groups in Table 4.3.5.

Table 4.3.5 **Comparison of changes in the estimated probability of EBP from pre- to post-intervention for CAL-DS group versus others including case-mix**

Groups comparison	P value
CAL-DS V Non-intervention	0.5
CAL-DS V A&F	0.2
CAL-DS V A&F + CAL-DS	0.3

4.4 Section 4

4.4.1 Descriptive analyses of patients' characteristics:

Specific pathologies related to third molar teeth are considered as patients' characteristics. This section deals with the various characteristics of patients presenting at pre- and post-intervention stages.

4.4.1.1 Assessment of third molar problems using the patient as the unit of analysis

The SIGN guideline (SIGN 43, 2000) makes explicit recommendations on characteristic conditions such as pericoronitis, caries, untreatable pulp pathology and periodontal disease that may indicate the removal of the affected third molar tooth. The frequency of pathologies recorded in pre- and post-intervention was examined to evaluate the impact of the interventions on patient management with particular reference to these pathologies (Table 4.4.1).

The overall number of pathologies recorded for all the patients (N=670) in the study at pre- and post-intervention stages combined is presented in Table 4.4.1. Note that the total number of pathologies in each column does not add up to the total number of patients since some patients presented with multiple pathology. For clarity the total number of patients with multiple pathologies is presented in Table 4.4.2.

Table 4.4.1 Pre- versus post-intervention comparison of proportion of different diagnosed pathologies

Common Diagnosis	Pre-intervention	Post-intervention	Total	Pearson Chi-square
	N =244 n(%)	N =426 n (%)	N=670	*P-value df=1
Pericoronitis	152 (62)	228 (54)	380 (57)	0.03
Caries	72 (30)	136 (32)	208 (31)	0.6
Pulp pathology	30 (12)	47 (11)	77 (11.5)	
Periodontal disease	2 (1)	9 (2)	11 (2)	
Orthodontic / imbrication of lower incisors	1 (0)	5 (1)	6 (1)	
Overeruption	7 (3)	3 (1)	10 (1)	

*P values calculated for pre and post-intervention comparison. N=total number of patients, n= number of patient with a pathology

Table 4.4.2 Overall number of patients in whole study with co-existing pathologies

Shared pathologies	Overall number of patients N=670
Pericoronitis + pulpal pathology	2
Pericoronitis + caries	8
Caries + pulpal pathology	56

Total number of patients in the study =N

The most frequent pathologies diagnosed were pericoronitis and caries (Table 4.4.1). Patients were more likely to have pericoronitis pre-intervention (62%) than post-intervention (54%) (P=0.03). However, no significant statistical difference was observed in patients with caries at two stages of the study (P=0.6) (Table 4.4.1).

Examining the number of these common pathologies across the groups, it appears that the proportion of patients with pericoronitis in the CAL-DS group shows a drop from 77% (36) to 58% (52) in the pre- to post-intervention (P=0.04) (Table 4.4.3).

Pre-intervention, a relatively smaller number of caries cases (9 cases (18%)) were treated by dentists in the CAL-DS group in comparison with the other intervention groups. This number increased to 22 (24%) cases after intervention (P=0.01) (Table 4.4.4).

Table 4.4.3 The number of pericoronitis cases in different intervention groups at pre- and post-intervention

Groups	Pre- intervention	Post- intervention	Pearson Chi-square
	n/N (%)	n/N (%)	P Value
Non-intervention/control	36/65 (55)	52/94 (55)	0.9
A&F	47/72 (65)	91/165 (55)	0.2
CAL-DS	36/47 (77)	52/90 (58)	0.04
A&F + CAL-DS	33/60 (55)	33/77 (43)	0.2
Total	152/244 (62)	228/426 (53)	

N=total number of cases pre- or post-intervention, n=number of patient with pericoronitis per group

Table 4.4.4 The number of caries cases in different intervention groups at pre- and post-intervention

Groups	Pre- intervention	Post- intervention	Pearson Chi-square
	n/N (%)	n/N (%)	P Value
Non-intervention	24/65(37)	32/94(34)	0.8
A&F	21/72(29)	41/165(25)	0.6
CAL-DS	9/47(18)	22/90(25)	0.04
A&F + CAL-DS	18/60(30)	41/77(53)	0.01
Total	72/244 (30)	136/426(32)	

N=total number of cases pre- or post-intervention, n=number of patient with caries per group

Given that pericoronitis and caries were found to be common pathologies, they were subjected to further assessment.

4.4.1.1.a Recurrent status of pericoronitis

The SIGN guideline (SIGN 43, 2000) make a recommendation on the recurrent status of infections such as pericoronitis. It is clear if there have been recurrent acute attacks of infection then the early removal of affected teeth is encouraged. In addition, the SIGN guideline (SIGN 43, 2000) advocates that the removal of symptomatic third molars should be considered where there have been one or more episodes of infection. The recorded recurrent status of the pericoronitis was examined in this section to assess the level of concordance with the guideline.

For 30 patients, there was no record of the recurrent status of pericoronitis in their dental notes (21 at pre-intervention and 9 at post-intervention). For those who had a record, 51% experienced a recurrent incidence of pericoronitis, compared with 49% who had a single incident pre-intervention (Table 4.4.5).

Post-intervention, more patients (58%) who visited their dentist had experienced only one episode of pericoronitis, compared to 43% who had had recurrent episodes (Table 4.4.5). The pre- and post-intervention comparisons of the recurrent status of pericoronitis cases were not significant (P=0.2).

Table 4.4.5 Pre versus post-intervention comparison of recurrent status of pericoronitis

Recurrent status	Pre- intervention	Post- intervention	Pearson Chi-square
	N=131 n (%)	N=219 n (%)	**P value df=1**
Single	64 (49)	124 (58)	0.2
Recurrent	67 (51)	95 (43)	

N= total number of patients, n=number of patients with a record of their recurrent status of pericoronitis

The recurrent status of the pericoronitis cases were compared with guideline adherence (EBP). It was found that 68% of treatments of all the patients who had a recurrent pericoronitis complied with the guideline, in comparison with 65% of the treatments of patients with a single episode. This overall comparison was not statistically significant (P=0.7) (Table 4.4.6).

Pre -intervention, the treatment of 51% of the patients who had a recurrent pericoronitis complied with the guidelines (EBP) and this reduced to 45% in the post-intervention. The treatments of 49% of the patients who had a single episode of pericoronitis in the pre-intervention complied with guidelines. This was increased to 55% at post-intervention (Table 4.4.7). However, neither of these changes was statistically significant (P=0.4) (Table 4.4.7).

Table 4.4.6 **Comparison of recurrent status of pericoronitis with guideline adherence (EBP)**

Pericoronitis	Recurrent Status N=350		Pearson Chi-square
Guideline adherence (EBP)	**Multiple Episodes N= 162 n (%)**	**Single Episode N=188 n (%)**	**P value** df=1
Yes	110 (68)	123 (65)	0.7
No	52 (32)	65 (35)	

N=total number of patients, n=number of patients in each category

Table 4.4.7 **Pre- versus post-intervention comparison of adhered cases with guideline (EBP) for recurrent status of pericoronitis**

Pericoronitis		Pre-intervention	Post-intervention	Pearson Chi-square
Guideline adherence (EBP)	**Recurrent Status**	**N=131 n (%)**	**N=219 n (%)**	**P value** df=1
Yes N=233	**Multiple Episodes**	45/88(51)	65/145 (45)	0.4
	Single Episode	43/88(49)	80/145 (55)	
No N=117	**Multiple Episodes**	22/43(51)	30/74(40)	0.3
	Single Episode	21/43(49)	44/74(60)	

Hierarchical log linear analysis was used to examine the effect of time, i.e. the comparison of pre- with post-intervention and EBP (adherence with the guidelines) on the recurrent status of pericoronitis and their interaction (Table 4.4.8).

Log linear analysis confirmed the above findings (Table 4.4.5, 4.4.6, 4.4.7) and revealed that time and EBP are not significantly associated with the recurrent status of pericoronitis (P=0.99, P=0.35 respectively). Their three-way interaction was not significant (P=0.99) either (Table 4.4.8). This confirms that there is no three way association between "pre- and post-intervention comparisons of "adhered cases with the guideline" and "the recurrent status of pericoronitis" (Table 4.4.7, 4.4.8).

This implies that frequency of a recurrent or a single episode of pericoronitis does not change significantly post-intervention (P=0.99). There is no statistically significant change in the compliance of the treatment (EBP) of patients presenting with recurrent pericoronitis in comparison with the single episode presentations (P=0.35). The EBP treatment of patients presenting with recurrent or single episodes of pericoronitis did not change significantly following the intervention phase (P=0.99).

Table 4.4.8 Test of association (Hierarchical Log linear) of recurrent status of pericoronitis with Time (i.e. pre- to post-intervention) and EBP

	DF	P value
Recurrent status* EBP	1	0.99
Recurrent status* Time	1	0.35
Time* EBP * recurrent status	1	0.99

Time= the pre- and post-intervention comparison

Post-intervention, the 6% increase in the compliance of treatments of single episodes of infection and the 6% reduction in the compliance of treatment of multiple episodes of infection were not statistically significant (P=0.4). This seems to suggest that the recurrent status of pericoronitis was not a strong indicator of EBP-complaint treatment.

4.4.1.2 Assessment of the overall treatments for third molar teeth

Overall, the treatments recommended by the SIGN Guideline (SIGN 43, 2000) for third molar teeth fall into four broad categories: referral, extraction, observation and restoration. The observation category included conservative treatments of third molar teeth such as mouthwash, local irrigation and the prescription of antibiotics (see Chapter 3, Section 3.6.2.5).

The overall treatments provided by participating dentists for third molars were assessed before and after intervention. Among all the different treatment categories, the most frequent treatments provided for patients with third molar pathologies were extraction with 37% pre-intervention and observation with 40% post-intervention (Table 4.4.9).

The pre- and post-intervention comparison of treatments recorded for third molar problems revealed an overall statistically significant change in the proportion of different treatment categories (P=0.001), (Table 4.4.9).

Examining individual treatment categories, the pre and post-intervention comparison of referral and extraction treatment categories showed that the relative frequencies of referral and extraction cases remained almost the same (P=0.99), (Table 4.4.10). However, the pre and post-intervention comparison of the observation with restoration treatment categories showed that there was a statistically significant change in the relative frequencies of these treatments categories from pre- to post-intervention (P=0.001), (Table 4.4.11).

Additionally, for the purpose of analysis, the treatments categories were arbitrarily divided into two broader groups: "proactive" and "reactive" (see Chapter 3. Section 3.6.2.5). Pre- and post-intervention comparison of these two groups revealed that proactive (extraction and referral) treatments decreased compared to reactive (restoration and observation) treatments (P=0.001), (Table 4.4.12).

Table 4.4.9 Pre-versus post-intervention comparison of overall treatments

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=244** **n (%)**	**N=426** **n (%)**	**P value** df=1
Refer	47 (19)	63 (15)	0.001
Extract	89 (37)	116 (27)	
Observation	86 (35)	169(40)	
Restore	22 (9)	78(18)	

N=total number of patients, n=number of patients in each treatment category

Table 4.4.10 Pre-versus post-intervention comparison of overall referral and extraction treatments

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=136** **n (%)**	**N=179** **n (%)**	**P value** df=1
Referral	47 (35)	63 (35)	0.99
Extraction	89 (65)	116 (65)	

N=total number of patients, n=number of patients in each treatment category

Table 4.4.11 Pre-versus post-intervention comparison of overall observation and restoration treatments

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=108** **n (%)**	**N=247** **n (%)**	**P value** df=1
Observations	86 (80)	169 (68)	0.04
Restorations	22 (20)	78 (32)	

N=total number of patients, n=number of patients in each treatment category

Table 4.4.12 Pre-versus post-intervention comparison of overall treatments "proactive" and "reactive"

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=244** **n (%)**	**N=426** **n (%)**	***P value** df=1
Proactive	136 (56)	179 (42)	0.001
Reactive	108 (44)	247 (58)	

N=total number of patients, n=number of patients in each treatment category

Examination of adherence to the guideline (EBP) revealed that a high proportion of all treatment categories have followed the EBP at both phases of the study, apart from the "observation" category. However, the overall effect, i.e. changes in EBP for all treatment groups at post-intervention were statistically significant (P=0.0001) (Table 4.4.13).

Pre- and post-intervention comparison of each treatment category was carried out. In the referral treatment category, the proportion of the EBP reduced by 6% post-intervention. This reduction was statistically non-significant (P=0.4). In the observation treatment category, the proportion of the EBP rose by 18% in post-intervention. This change was statistically significant (P=0.008), (Table 4.4.13).

Table 4.4.13 **Pre- versus post-intervention comparison of each treatment category (and also overall treatments) which adhered to guideline (EBP)**

	Pre-intervention	Post-intervention	Pearson Chi-square	
			P value	
Agreed treatments with guidelines (EBP)	**n/N (%)**	**n/N (%)**	(break down) for each treatment category df=1	(combine) for overall treatment effect df=3
Referral	46/47 (98)	58/63 (92)	0.4	0.0001
Extraction	83/89 (93)	109/116 (94)	0.8	
Observation	30/86 (35)	90/169 (53)	0.008	
Restoration	22/22 (100)	77/78 (99)	0.6	

n=number of EBP treatment cases in each treatment category, N= total number of cases in each treatment category

4.4.1.2a Extraction

A comparison of the number of cases in the extraction category with other treatment categories, *i.e.* referral, observation and restoration was collectively carried out (Table 4.4.14). A statistically significant reduction in the overall number of patients treated with extraction was detected between the pre- (37%) and post-intervention (27%) phase of this study, (P=0.02) (Table 4.4.14) and this rate did not differ significantly for different groups (P>0.05) (Table 4.4.15). Hierarchical Log Linear analysis was used to examine the effect of time, *i.e.* comparison of pre- to post-intervention, on the extraction and intervention groups and their interaction (Table 4.4.16). Log Linear analysis confirmed the above findings (Table 4.4.14, 4.4.15) and revealed that time is significantly associated with extraction (P=0.02), (Table 4.4.14, 4.4.16). However, the intervention groups are not significantly associated with extraction (P=0.26), (Table 4.4.15, 4.4.16). The three-way interaction was not significant either (P=0.99), (Table 4.4.16) in that there was no association observed between the intervention groups and the extraction rate from the pre- to post-intervention.

Table 4.4.14 Pre- versus post-intervention comparison of overall treatment of extraction with other treatments

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=244** **n (%)**	**N=426** **n (%)**	**P value** df=1
Extraction	89 (37)	116 (27)	0.02
Other treatment	155 (63)	310 (73)	

N=total number of patients, n=number of patients in each treatment category

Table 4.4.15 Pre- versus post-intervention comparison of "extraction" versus "other treatments" for different intervention groups

		Pre-intervention	Post-intervention	Pearson Chi-square
Group	**Treatments**	**n/N (%)**	**n/N (%)**	***P value df=1**
Non-int/ Control	Extract	28 (43)	30 (32)	0.2
	Other treatments	37(57)	64 (68)	
A&F	Extract	23 (32)	39 (27)	0.2
	Other treatments	49 (68)	126 (76)	
CAL-DS	Extract	16 (34)	25 (28)	0.6
	Other treatments	31 (66)	65 (72)	
CAL-DS + A&F	Extract	22 (37)	22 (29)	0.4
	Other treatments	38 (63)	55 (71)	

N=overall number of treatments in each group, n=number of each treatment for each group

Table 4.4.16 Test of association (Hierarchical Log linear) of extraction with Time (i.e. pre- to post-intervention) and group

	DF	P value
Time* extraction	1	0.02
Group* extraction	3	0.26
Time* group* extraction	3	0.99

Time= the pre- and post-intervention comparison

At post-intervention, a statistically significant reduction in the overall number of patients treated with extraction was observed. This is in contrast to no significant change in the rate of extractions with the different intervention groups.

4.4.1.3 Assessment of treatments for the common pathologies in this study

Pericoronitis and caries were the most common pathologies recorded in this trial. The assessments of the treatments recorded for each of these pathologies are outlined below:

4.4.1.3a Pericoronitis treatments

The treatments for pericoronitis were assessed and tabulated (Table 4.4.17). The predominant treatment category recorded for the patients with pericoronitis was "observation", with 47% and 59% respectively in pre- and post-intervention (Table 4.4.17). There was insufficient data for the "restoration" group (2 cases pre- and 1 case post-intervention) so this was excluded from the analysis. The pre- and post-intervention comparison of the treatment categories for pericoronitis showed that post-intervention there were reductions in the percentages of referrals and extractions treatment categories (21% and 20% respectively) in comparison with the pre-intervention referrals and extractions (28% and 24% respectively) (Table 4.4.17). However, these overall changes were not statistically significant (P=0.08) (Table 4.4.17).

On examining the individual treatment categories, the pre- and post-intervention comparison of the relative proportion of the "referral" and the "extraction" groups revealed that the changes in these treatment categories were not statistically significant (P=0.8) (Table 4.4.18).

Pre- and post-intervention comparison of the two broad categories of treatments (i.e. "reactive" and "proactive") for pericoronitis revealed that post-intervention the

proportion of the "reactive" (observation) treatment cases was higher than "proactive" (referral and extraction) treatment cases (P=0.03) (Table 4.4.19).

The different treatment recorded for pericoronitis in different intervention groups was also examined. Pre- and post-intervention comparison of different groups distinguished the CAL-DS group from the others in terms of an overall statistically significant change in the percentage of different treatment categories in this group (P=0.01) (Table 4.4.20). The Hierarchical Log Linear analysis was subsequently used to examine the effect of different groups and of time i.e. pre- and post-intervention period on the treatment recorded for pericoronitis and their interaction. This method showed that different intervention groups were not significantly associated with time or treatment (P=0.2, P=0.2 respectively) (Table 4.4.21). Time was significantly associated with treatment of pericoronitis though (P=0.01) (Table 4.4.21). The effect of the groups was that the independence of time and treatment interaction and the three-way interaction was not significant. There was no significant association observed between the groups and the treatment of pericoronitis across the pre- and post-intervention phases (P= 0.2) (Table 4.4.21), that is the effect seen in the CAL-DS group (Table 4.4.20) is nothing but the statistically significant change of overall treatment seen from pre- to post-intervention and not the true effect of CAL-DS intervention.

Table 4.4.17 **Pre- versus post-intervention comparison of treatments recorded for pericoronitis (excluding restoration cases)**

Pericoronitis	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=150** **n (%)**	**N=227** **n (%)**	**P value** df=2
Refer	42 (28)	48 (21)	0.08
Extract	37 (24)	45 (20)	
Observation	71 (47)	134 (59)	

N=overall number of treatments, n=number of each treatment category

Table 4.4.18 **Pre- versus post-intervention comparison of referral and extraction treatments for pericoronitis**

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=79** **n (%)**	**N=93** **n (%)**	**P value** df=1
Referral	42 (53)	48 (52)	0.8
Extraction	37 (47)	45 (48)	

N=overall number of treatments, n=number of each treatment category

Table 4.4.19 **Pre- versus post-intervention comparison of reactive (excluding restoration cases) and proactive treatments (referral and extraction) for pericoronitis**

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=150** **n (%)**	**N=227** **n (%)**	**P value** df=1
Observation (reactive)	71 (47)	134 (59)	0.03
Proactive	79 (53)	93 (41)	

N=overall number of treatments, n=number of each treatment category

Table 4.4.20 **Pre- versus post-intervention comparison of treatments for pericoronitis (excluding restoration cases) for each intervention group**

	Pericoronitis	Pre-intervention	Post-intervention	Pearson Chi-square
Group	**Treatments**	**N=150 n (%)**	**N=227 n (%)**	**P value**
Non-	Referral	8 (22)	10(19)	0.6
	Extraction	13(36)	15(29)	
	Observation	15(42)	27(52)	
A&F	Referral	8(17)	22(24)	0.5
	Extraction	10(21)	14(15)	
	Observation	29(62)	55(60)	
CAL-DS	Referral	14(41)	9(18)	0.01
	Extraction	9(27)	9(18)	
	Observation	11(32)	33(64)	
CAL-DS + A&F	Referral	12(36)	7(21)	0.4
	Extraction	5(15)	7(21)	
	Observation	16(49)	19(58)	

N=overall number of treatments, n=number of each treatment category

Table 4.4.21 **Tests of association (Hierarchical Log Linear) of treatments for pericoronitis with Time (i.e. pre- to post-intervention) and groups**

	DF	P value
Time* Treatment	2	0.01
Groups* Time	3	0.2
Groups * Treatment	6	0.2
Groups* Time* Treatment	6	0.2

Time= the pre- and post-intervention comparison

The comparison of the degree of the EBP for different treatment categories revealed that all treatments significantly complied with the guidelines (EBP) whether they were provided pre- or post-intervention (P=0.0001). In other words, the high percentages of treatments followed EBP in both pre- and post-intervention (Table 4.4.22). However, the pre- and post-intervention comparison of EBP treatments (*i.e.* treatments which complied with the guideline) for pericoronitis demonstrated that there was no

statistically significant change in the overall rate of EBP treatment from pre- to post-intervention (P=0.7), (Table 4.4.23).

Table 4.4.22 EBP (adherence to guidelines) comparison of different treatments (excluding restoration cases) for pericoronitis

Pericoronitis			Treatments compliance with guidelines (EBP)		Pearson Chi-square
Intervention		**Treatments**	**Yes**	**No**	**P value** df=2
Pre-	**n/N (%)**	**Refer**	41/42 (98)	1/48 (2)	0.0001
		Extract	35/37 (95)	2/48 (5)	
		Observation	26/71 (37)	45/48 (63)	
Post-	**n/N (%)**	**Refer**	48/48(100)	0(0)	0.0001
		Extract	45/45 (100)	0 (0)	
		Observation	58/134 (43)	76/134 (57)	

N=Total number of each treatment category (excluding restoration), n=number of EBP (or non-EBP) treatment for pericoronitis in each treatment category

Table 4.4.23 Pre- versus post-intervention comparison of EBP treatment for pericoronitis

Agreements with guidelines (EBP)		Pre-intervention	Post-intervention	Pearson Chi-square P value df=1
Yes	**N=255, n (%)**	104(41)	151(59)	0.7
No	**N=125, n (%)**	48 (39)	77(61)	

N=total number in each category, n=number of cases in each category

Investigating the rate of the EBP treatment for each treatment category for pericoronitis shows that pre-intervention, a high percentage of referral (98%) and extraction (95%) cases followed the EBP when compared to the observed cases (37%). A similar pattern occurred at post-intervention, with 100% of referrals and extraction cases and 43% of observed cases following the EBP. However, comparing pre- and post-intervention, the

rates of the EBP did not significantly change in either treatment category (P>0.05) (Table 4.4.24).

Table 4.4.24 Pre- versus post-intervention comparison of EBP treatment for each treatment category for pericoronitis

	Pre-intervention	Post-intervention	Pearson Chi-square
Agreed treatments with guidelines (EBP)	**n/N(%)**	**n/N (%)**	**P value** df=1
Refer	41/42 (98)	48/48(100)	0.5
Extract	35/37 (95)	45/45 (100)	0.2
Observation	26/71 (37)	58/134 (43)	0.4

N=Total number in each treatment category, n=EBP treatments in each treatment category

Log Linear analysis confirmed the above findings (Table 4.4.22, 4.4.23, 4.4.24). It revealed that treatments of pericoronitis were not equally frequent between pre- and post-intervention (as already seen in Table 4.4.19), (P=0.03) (Table 4.4.25). Treatments of pericoronitis differed in compliance with the guidelines (EBP) (P=0.0001). Compliance of treatment with the guideline (EBP) did not differ with time, i.e. comparing pre- and post-intervention (P=0.2). The effects of time and treatments on compliance with the guidelines (EBP) were independent and the three-way interaction between these elements was not significant (P=0.3) (Table 4.4.25).

Table 4.4.25 Test of association (Hierarchical Log Linear) of treatments for pericoronitis with Time (*i.e.* pre- to post-intervention) and EBP

	DF	P value
Time * treatment	2	0.03
Treatment * EBP	2	0.0001
Time * EBP	1	0.2
Time* treatment * EBP	2	0.3

Time= the pre- and post-intervention comparison

The results seem to suggest that following the intervention, dentists carried out fewer extractions or referrals for patients presenting with pericoronitis and kept patients under observation more often. A large number of treatments provided followed EBP whether they were observed at the pre- or post-intervention phases. There was no difference in rate of EBP-treatment between pre- and post-intervention.

4.4.1.3b Caries treatment

Investigating the treatments for patients with third molar caries shows that the predominant treatments for caries were extractions (pre- 61% and post- intervention 40%) followed by restoration (pre- 31% and post intervention 56%). Comparing pre- and post-intervention, these changes noted above, i.e. fall in treatment of extraction by 19% and rise in restoration treatment by 25% from pre- to the post-intervention, were statistically significant (P=0.001) (Table 4.4.26). (Similar comparisons for the "observation" and the "referral" categories were not carried out because of the low numbers of patients in each category (i.e. for observation: 1 case pre- and 1 case post-intervention and for referral: 5 cases pre- and 4 cases post-intervention)).

Table 4.4.26 Pre- versus post-intervention comparison of treatments recorded for caries (excluding referral and observed cases)

	Pre-intervention	Post-intervention	Pearson Chi-square
Treatments	**N=66** **n (%)**	**N=131** **n (%)**	**P value** df=3
Extraction	44 (67)	54(41)	0.001
Restoration	22 (33)	77(59)	

N=Total number of treatments, n=number of each treatment

Pre- and post-intervention comparisons of the EBP treatment shows that all treatments for caries appear to have followed EBP whether they were observed at the pre- or post-intervention (Table 4.4.27).

Table 4.4.27 Pre- versus post-intervention comparison for each treatment category for caries which followed EBP (adhered to guideline)

	Pre-intervention	Post-intervention	Pearson Chi-square
Agreed treatments with guidelines	**n/N (%)**	**n/N (%)**	**P value** df=1
Extraction	43/44 (98)	54/54 (100)	0.45
Restoration	22/22 (100)	76/77 (99)	0.99

N=Total number of treatments in each treatment category, n=number of EBP treatments in each treatment category, (excluding referral and observed cases)

Log Linear analysis confirmed the above findings (table 4.4.26, 4.4.27). It revealed that treatment for caries was not equally frequent between pre- and post-intervention (P=0.003), (as already shown in Table 4.4.26) and the treatments of caries did not differ in compliance with the guideline (EBP) (P=0.99). Compliance of treatment did not differ from pre- to post-intervention (EBP) (P=0.6). The effects of time (i.e. from pre- to post-intervention) and treatment on EBP were independent. The three-way interaction

of time (i.e. from pre- to post-intervention) EBP and treatment was not significant (P=0.99) and there is seems to be no association between them (Table 4.4.28).

Table 4.4.28 Test of association (Hierarchical Log linear) of treatments for caries with Time (*i.e.* pre- to post-intervention) and EBP

Caries	DF	P value
Treatment * time	2	0.003
Treatment * EBP	2	0.99
Time*EBP	1	0.6
Time* EBP * treatment	2	0.5

Time= the pre- and post-intervention comparison

The results seem to suggest that dentists carried out a higher number of restorations and a lower number of extractions for patients with caries at post-intervention. All treatments provided for caries appear to have followed EBP irrespective of the period of the study.

The SIGN guideline (SIGN 43, 2000) has recommendations in respect to radiographic examinations and recording medical history when removal of third molar teeth is considered. In this study, these two recommendations are examined in the next two sections.

4.4.1.4 Availability of a radiograph and agreement with guidelines

In this section, the availability of radiographs and the rate of the adherence of treatments to the guideline were examined. Of 670 patients, 12% (n=78) had a radiograph available at the time of data collection. Of those (n=78) who had a radiograph, 10% (n=25) were

taken at the pre- and 12% (n=53) at the post-intervention. Ninety five percent (n=74) of these patients' treatment complied with the EBP (P=0.001) (Table 4.4.29). This level of compliance with the guidelines (EBP) corresponded to the respective values of 92% (n=23) at the pre- and 96% (n=51) at the post-intervention. This increase in the percentages of EBP treatments when a radiograph was available was not statistically significant (P=0.6)(Table 4.4.30).

Table 4.4.29 Guideline adherence/EBP when a radiograph was available

Radiographs	Adherence with guideline /EBP		N=670	Fisher exact test
Availability	Yes	No	Total (%)	P value df=1
	n (%)	n (%)		
No	441 (74)	151 (26)	592 (88)	0.001
Yes	74 (95)	4 (5)	78 (12)	

N=Total number of patients in each row, n=number of EBP treatments

Table 4.4.30 Pre- versus post-intervention comparison of EBP (adhered with guidelines) treatment when a radiograph was available

Treatments	Available radiographs		N=78	Fisher exact test
Agreed with guidelines	Pre-	Post-		P value df=1
	N=25 n (%)	N=53 n (%)	N=78 n (%)	
	23(92)	51(96)	74(95)	0.6

The results seem to suggest that when a radiographic examination was carried out by GDPs, a high proportion of the patient's treatments followed EBP in both the pre- and post- intervention.

4.4.1.5 Medical History

In this section, recorded medical history was examined and a comparison of the presence of a recorded medical history between pre- and post-intervention was carried out. Sixty-nine percent of all patients had a medical history recorded in their dental notes at the post-intervention compared with 47% at the pre-intervention stage (Table 4.4.31). This difference is statistically significant ($P < 0.01$).

The assessment of the recorded medical history across all the intervention groups revealed that overall a higher proportion of patients in the audit and feedback group had their medical history recorded in their dental notes (Table 4.4.32). Comparing the pre- and post- intervention, an assessment of the recorded medical history in the different groups showed that in all groups more patients had a medical history recorded in their dental notes at the post-intervention in comparison with the pre-intervention. These differences were only statistically significant for the "A&F" and the "CAL-DS+A&F" (Table 4.4.33), (this is not true effect for these two groups as shown in Table 4.4.34).

The Hierarchical Log Linear model was employed to examine the effect of time (*i.e.* comparison of pre- with post-intervention) on the recorded medical history of different intervention groups and their interaction (Table 4.4.34). It was found that time is significantly associated with recorded medical history ($P=0.0001$) and a higher proportion of patients had a medical history recorded in their dental notes at post-intervention. This two-way interaction was independent of the groups.

In addition, the groups are significantly associated with recorded medical history (P=0.0001) and overall a higher proportion of patients in the A&F group had their medical histories recorded in their dental notes. This two-way interaction is independent of time and there is no association between the effects of intervention groups on recording a medical history in dental notes across pre- and post-intervention phases (Table 4.4.34). The recording of medical history did not differ significantly in different intervention groups comparing the pre- and post-intervention (Table 4.4.34).

Table 4.4.31 Pre- versus post-intervention comparison of recorded medical history

Medical History	Pre-intervention	Post-intervention	Pearson Chi-square
Recorded	**N=244** **n (%)**	**N=426** **n (%)**	**P value** df=1
Yes	115 (47)	292 (69)	0.0001
No	129 (53)	134 (31)	

N=total number of patients, n=number of patient with/without a radiograph

Table 4.4.32 Recorded medical history for all four groups

	Medical history			Pearson Chi-square,
Groups	**Yes** n (%)	**No** **n (%)**	**Total** **N=670**	**P value** df=3
Non-intervention/control	93 (59)	66 (41)	159 (24)	0.0001
A&F	173(73)	64 (27)	237 (36)	
CAL-DS	64(47)	73 (53)	137 (20)	
CAL-DS & A&F	77(56)	60 (44)	137 (20)	

N=total number of patients, n=number of patient with/without a radiograph in each intervention group

Table 4.4.33 Pre- versus post-intervention comparison of a recorded medical history in each intervention group

Groups	Pre-intervention n/N(%)	Post-intervention n/N (%)	Total n (%)	Pearson Chi-square P-value df=1
Non-intervention/control	32/65 (49)	61/94 (65)	93/159 (59)	0.07
A&F	40/72 (56)	133/165 (81)	173/237 (73)	0.0001
CAL-DS	18/47 (38)	46/90 (51)	64/137 (47)	0.2
CAL-DS & A&F	25/60 (42)	52/77 (68)	77/137 (56)	0.004

N=total number of patients in each intervention group, n=number of patient with a radiograph in each intervention group

Table 4.4.34 Tests of association (Hierarchical Log linear) of medical history with Time (*i.e.* pre- to post-intervention) and EBP

	DF	P value
Time* recorded medical history (RMH)	1	0.0001
RMH* Group	1	0.0001
Time*RMH*Group	3	0.42

Time=comparing pre- and post-intervention

Hence, overall a higher number of patients in the post-intervention period had a medical history recorded in their dental notes by their GDPs than patients in the pre-intervention.

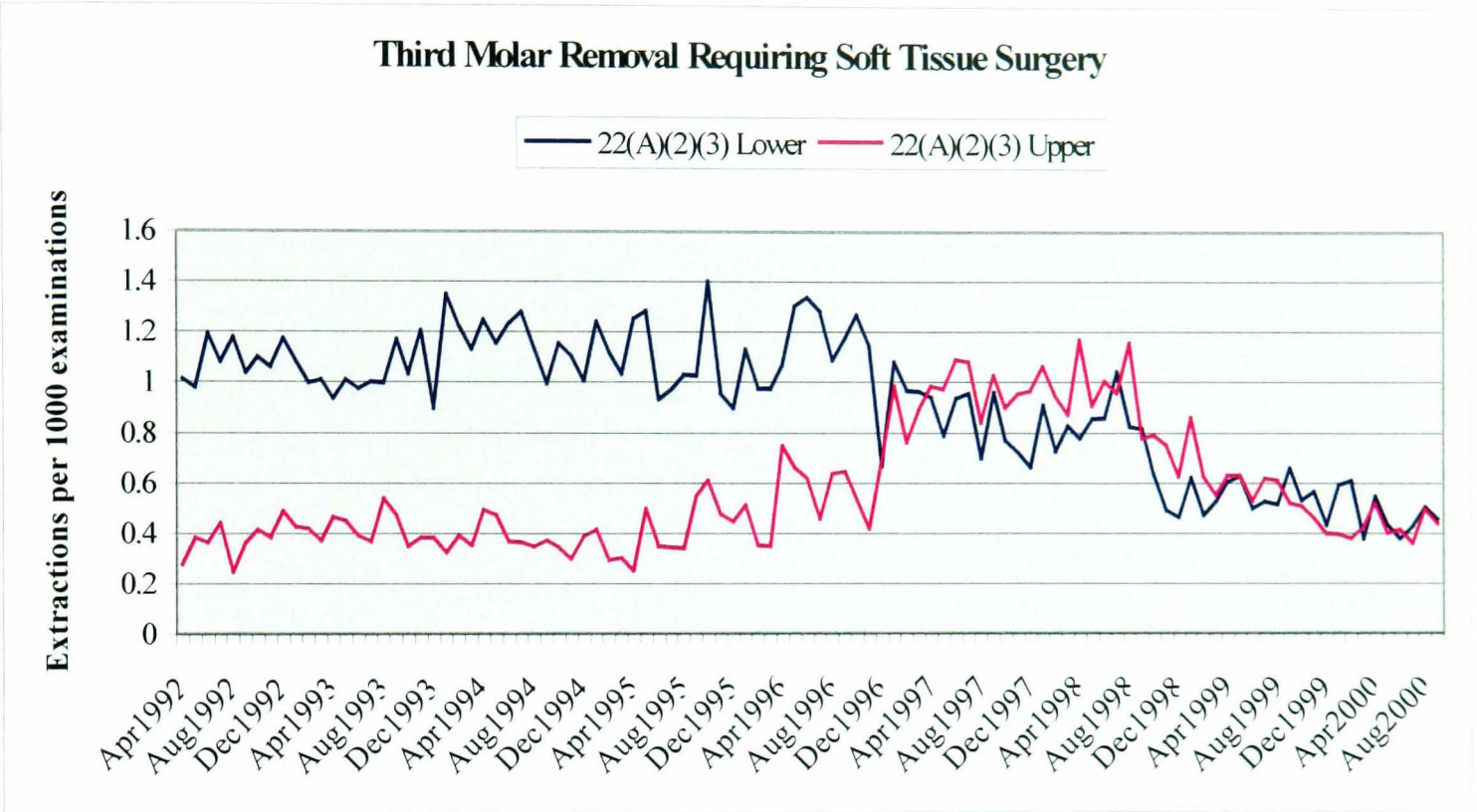
4.5 Section 5

4.5.1 Change in the rate of extraction of third molar teeth by Scottish dentists

In this section, removals of the third molar teeth by Scottish dentists were examined to assess the change in rate of removal of these teeth in a selected population of practitioners using data from NHS Scotland's Management Information and Dental Accounting System (MIDAS). The data show a reduction in both surgical and non-surgical wisdom tooth extractions took place in Scottish General Dental Practices before and during the experimental period (Graphs 4.5.1 and 4.5.2).

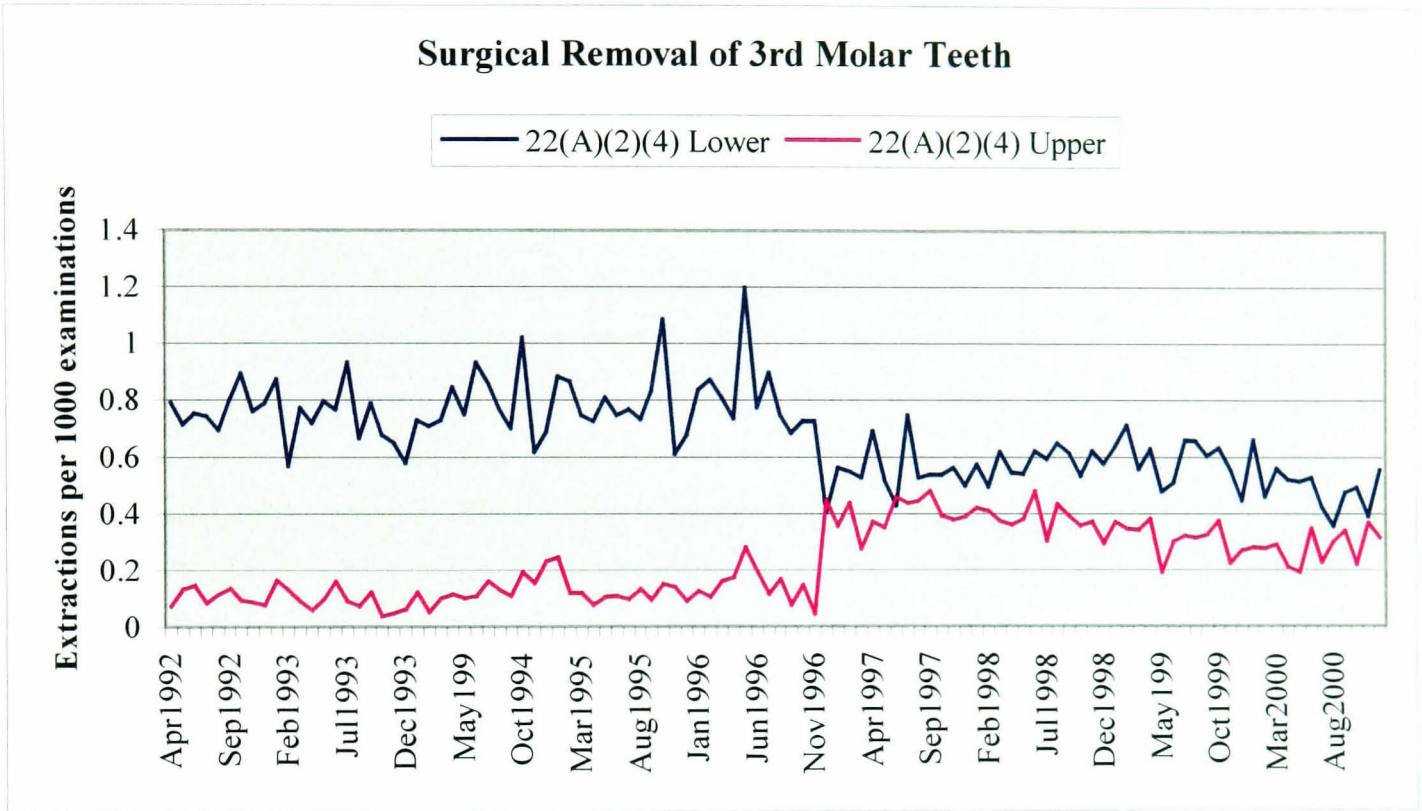
Graph 4.5.1

NHS Management Information and Dental Accounting System (MIDAS) data for third molar extraction between April 1992 and April 2000.



Graph 4.5.2

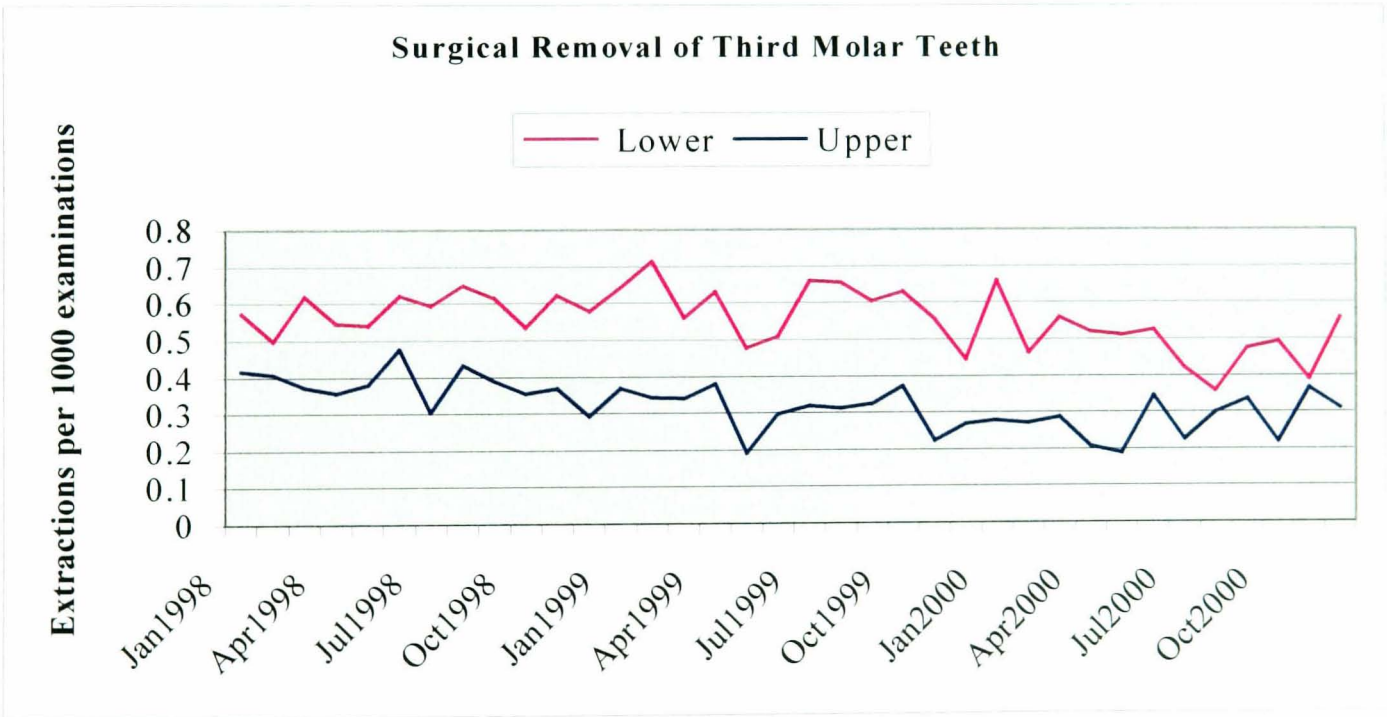
NHS Management Information and Dental Accounting System (MIDAS) data for surgical removal of third molar teeth between April 1992 and April 2000.



On examination of the data over a shorter period, the proportion of patients who had a third molar tooth removed by Scottish General Dental Practitioners over a 36-month period (January 1998 to January 2001) were compiled from MIDAS (Graph 4.5.3 and 4.5.4). We were hoping this would enable us to detect any seasonal changes in the proportion of attendees to these practices over the data collection periods. However, no discernible reduction in the proportion of these patients was detected (Graph 4.5.3 and 4.5.4).

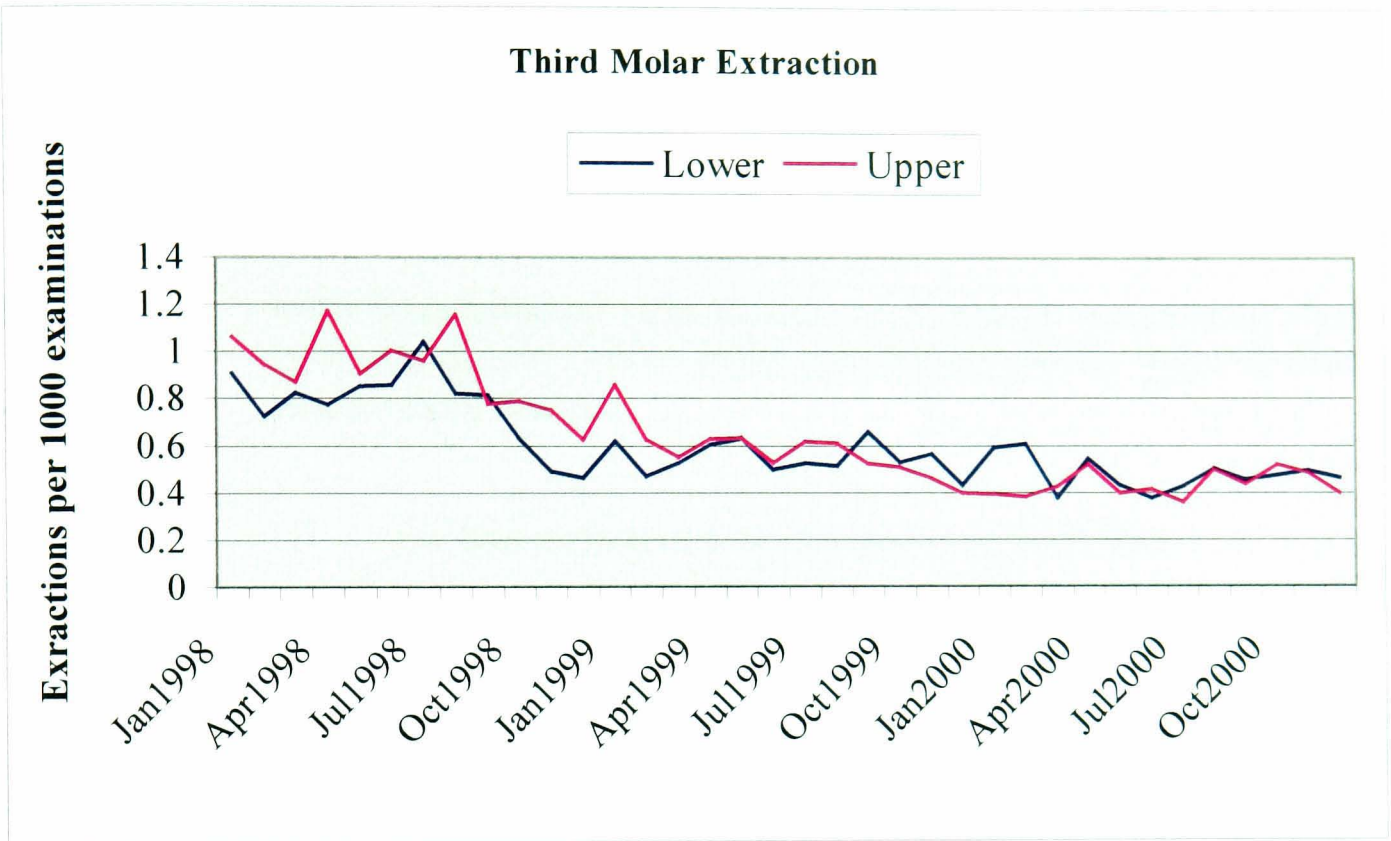
Graph 4.5.3

NHS Management Information and Dental Accounting System (MIDAS) data for surgical removal of third molar teeth between January 1998 and October 2000.



Graph 4.5.4

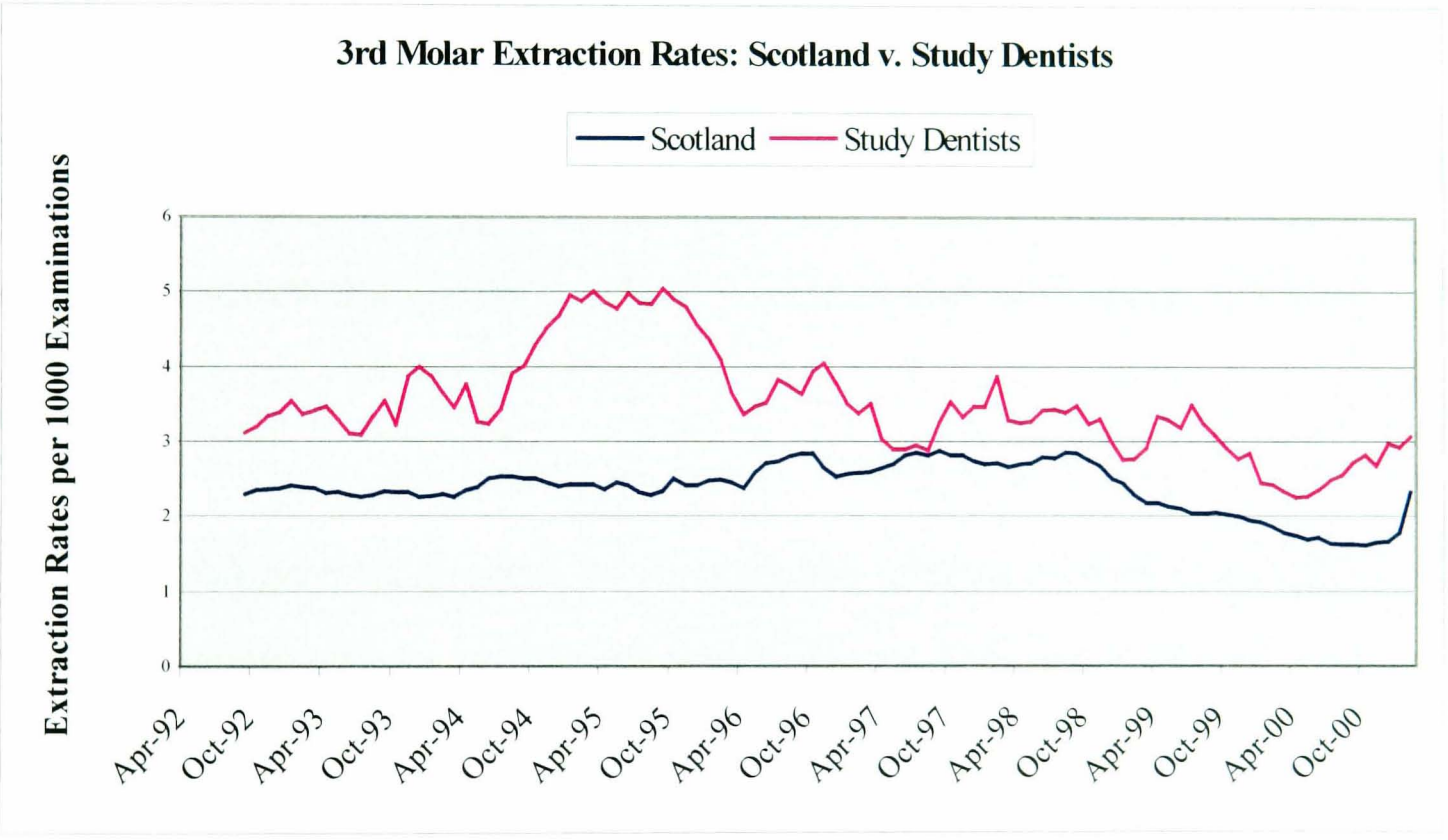
NHS Management Information and Dental Accounting System (MIDAS) data for third molar extraction between January 1998 and Oct 2000.



4.5.2 A comparison between the study population and others

A comparison of the third molar extraction rates of the study sample population of dentists with the other dentists in Scotland over a period of 8 years, using data from MIDAS (Graph 4.5.5) shows that the participating dentists removed more third molar teeth than other dentists in Scotland.

Graph 4.5.5 NHS Management Information and Dental Accounting System (MIDAS) data for third molar extraction rates for Scottish dentists versus study dentists



4.6 Section 6

4.6.1 Summary

Practice recruitment

Of 565 dentists invited to participate, at least one dentist from 63 practices volunteered to participate. Twelve of these practices withdrew before pre-intervention data. The 51 remaining dental practices were randomised into four groups. During the period between pre-intervention and post-intervention data collection four practices withdrew from the trial, leaving 47. The reasons given for withdrawal were too busy, not interested, moving practice, refurbishment of practice and a change of mind. There were no significant differences between the dentists who withdrew from the study and those who continued, in terms of their age (t=-1.34, P=0.2); gender (X^2=0.15, P=0.7); postgraduate qualifications (X^2=2.24, P=0.1) or their intervention group (X^2=4.24, P=0.2).

Of the 51 pre-intervention practices, 47 provided data. Of the 47 post-intervention practices 46 had relevant data. Of the dentists recruited, 23 (45%) attended one of the two postgraduate courses.

Pre-intervention, data were collected for 3342 (M=1885, F=1457) patients with a mean age of 21.7 years (SD 2.2, range 14-25.5 yrs) compared with 1935 (M= 880, F=1055) patients at the post-intervention stage with a mean age of 21.8 years (SD 2.1, range 16.6-25 yrs). The proportion of patients with a problem with their third molar teeth was 7% before and 22% after intervention. Female patients were more likely to have third molar problems than males in both the pre and post-intervention phases.

Outcome measure

The outcome measure for this study was the proportion of patients whose treatment followed the guideline, *i.e.* treatment which complied Evidence-Based Practice (EBP). Overall compliance with the guideline (EPB) at the pre-intervention stage was assessed to be 74% and increased to 78% at post-intervention. The adherence to the guideline (EBP) for different intervention groups showed a rise in compliance in all groups following the interventions; nevertheless this increase does not seem to be of statistical significance.

Effectiveness of the interventions

Comparison of pre-intervention data between the groups indicates that there was no apparent imbalance. The main effect of the intervention was examined at the cluster level and further at the patient level. Both analyses gave very similar results.

The weighted *t*-test for A&F versus no A&F was not statistically significant (P=0.62) neither was that between CAL versus no CAL (P=0.76).

From the multilevel analysis the odds ratio of compliance with guidelines for dentists who experienced A&F versus those who did not was 1.28 (95% CI 0.62 - 2.63) and the odds ratio was 0.84 (95% CI 0.88 - 1.74) for the CAL dentists versus no CAL. For neither was the difference statistically significant (Table 5).

The relative changes in EBP within the intervention groups from pre- to post-interventions were examined (sub-analysis). The probability of the EBP was estimated

from a regression equation and allowance was made or not in the model for patient characteristics (case-mix). The results showed the change in EBP within any group from pre- to post-intervention without a control for the case-mix was not statistically significant. Once a control for the case-mix was allowed for, a statistically significant impact of the CAL-DS intervention was observed. However, the intra-group changes i.e. the pre- and post- EBP differences within the groups, did not significantly differ across the different intervention groups.

Descriptive analysis results

Results showed that the most frequent pathologies diagnosed by participating dentists at both periods of the study were pericoronitis and caries. The overall treatment and the respective treatment categories for each of these pathologies were examined in more detail. The treatments were categorised into two broad groups of proactive (extraction and referral) and reactive (restoration and observation) treatments. Following the interventions, the overall treatments provided by participating dentists became more reactive (P=0.001). The rate of third molar extractions decreased for 16-24 years olds after the introduction of the SIGN Guideline. This reduction was statistically significant between the pre- and post-intervention phase of this study (37% to 27%, P=0.02) as the rate of extraction did not differ significantly between the groups.

Concurrent with the study data, data from the NHS management information and dental accounting system (MIDAS) shows a concomitant reduction in both surgical and non-surgical wisdom tooth extraction rates in Scottish general dental practices during the experimental period.

Examining the treatments provided for the patients presenting with pericoronitis revealed that dentists carried out less extractions or referrals and kept their patients more under observation at post-intervention. A large number of treatments provided followed EBP irrespective of the phase of the study. There was no difference in the compliance of the treatment with the guidelines (EBP) comparing pre- and post-intervention.

Similarly, investigating the treatments instituted for the patients presenting with third molar caries showed that the dentists carried out higher number of restorations and smaller numbers of extractions following any intervention. All treatments provided for caries appear to have followed the EBP irrespective of the period of the study.

The assessment of the overall recorded medical history considering both phases of the study together, across all the intervention groups revealed that overall a higher proportion of patients in the A&F group had their medical history recorded in their dental notes. Comparison of pre- and post-intervention revealed that overall a higher proportion of patients with a medical history were found in post-intervention than pre-intervention and this comparison was not statistically significant for any intervention group.

Conclusion

In conclusion, the evaluation of the EBP showed a high degree of compliance with the guideline (EBP) at the pre- and post-intervention. The change in the EBP compliance following the interventions was not statistically significant. The adherence to the guidelines (EBP) for different intervention groups showed a rise in compliance for all groups following the interventions; nevertheless this increase was not significant.

The results showed that the overall treatments after intervention changed significantly more in favour of the reactive (restoration, observation) treatment and less in favour of the proactive (extraction, referral) treatment for these two pathology categories i.e. caries and pericoronitis.

The main analyses did not highlight any statistically significant differences between the intervention groups and there was no evidence to suggest that the interventions, *i.e.* A&F and CAL-DS were more effective than the control/non-intervention group, *i.e.* mailing of guideline and a postgraduate continuing education courses (PGCE). However the sub-analysis suggests, when allowance was made for the confounding effect of case-mix (*i.e.* different patient characteristics), that there was a significant effect of the CAL-DS intervention on the implementation of these guidelines.

CHAPTER FIVE

5 Discussion

In order to interpret the results of the trial, initially the influence of the methodology of the study will be considered and this will be followed by discussions on various aspects of the findings and their implications. Finally, a conclusion will be drawn and recommendations will be made for future research.

5.1 Methodological considerations

This randomised controlled trial evaluated two dissemination and implementation strategies which were designed to enhance the adherence to the SIGN guideline (SIGN 43, 2000) for the appropriate management of third molar teeth by Scottish dentists. This study has been one of the first rigorous evaluations of intervention strategies in the primary dental care setting.

5.1.1 The design

Selecting an appropriate research design facilitates a robust scientific conclusion. This allows a researcher to attribute any changes observed to the principle being evaluated (Russell *et al* 1993).

In Chapter 2 (Section 4) of this thesis the choices of potential study designs for the evaluation of implementation strategies were presented. These reflect the informed choice of design in this study. A pragmatic randomised controlled trial with a 2x2 factorial design was employed as the most appropriate method of investigation for the

intervention strategies proposed. Randomised controlled trials have long been recognised as the most reliable form of research design to assess health care innovations (Sacket *et al* 1996, Prescott *et al* 1999). The principal reason for conducting a randomised controlled trial is that it makes causal inferences easier and although it is not perfect, it is superior to other designs as it rules out most threats to internal validity (Cook and Campbell 1979). The random allocation of subjects ensures that there is no bias in the group assignment and that subjects in each allocated group are comparable in both known and unknown confounding factors. Studies which evaluate the clinical interventions and fail to randomise subjects blindly tend to report greater effects (Schultz *et al* 1995).

In this study, participating dental practitioners were randomised through the computer generation of a random number sequence to four different intervention groups. This randomisation was carried out by a statistician independent of the research team who did not have any interest in the subject of the research and did not know the study participants. The lack of direct involvement eliminated the opportunity to influence the random allocation of the participants and minimised the introduction of any bias in the effect size.

The effects of interventions designed to change behaviour can be modified by factors such as the context in which the intervention is introduced, characteristics of the study population and the outcome measured (Grimshaw *et al* 1995). The design of our study provided an opportunity to compare the relative effectiveness of different interventions

and possibly balance any confounding factors associated with the setting, study population or outcomes.

This trial is a pragmatic randomised trial in which interventions were compared in a realistic setting, *i.e.* dental practice. The intention was that conclusions from this trial would be applicable in any other health care settings such as medicine, nursing, pharmacy and professions allied to medicine.

A 2x2 factorial design allowed for an efficient use of research funds by producing two separate designs for the price of one (Pocock 1983). Each subject had a one in four chance of being randomised into each intervention group. This design provided an opportunity to study and test for any probable interaction between the two interventions.

The control offered by the design when data were collected before (baseline) and after (post-) the intervention provided a means of assessing any secular or sudden changes during the study period. Baseline data gathering made it possible to evaluate current practice of the participating dentists prior to publication of the guideline as a baseline. This provided information about the effect of guideline dissemination on current practice as well as examining data-gathering methodology and piloting the data entry form.

5.1.2 Interventions

Research into the dissemination and implementation of guidelines has been generally conducted in medicine. A range of interventions have been shown to be effective in

changing medical practice in favour of published guidelines but consistently effective intervention strategies have not yet been identified.

While medicine and dentistry may well be similar in many aspects, there may also be considerable differences between these professions given their different organisational and funding structures (McGlone *et al* 2001). Therefore not everything applicable to medicine necessarily applies to dentistry. Equally the effect of implementation strategies may differ between the two professions. Nonetheless, it has been shown that our knowledge of altering practice in dentistry particularly as it relates to the effective utilisation of clinical guidelines is somewhat limited (McGlone *et al* 2001). However, there have been a few studies in dentistry which have evaluated the effectiveness of different interventions strategies (O'Brien *et al* 2000, Goodey *et al* 2000, Kay *et al* 2001).

Accordingly, the pressing need for undertaking research to identify the most effective strategies for the dissemination and implementation of guidelines in primary dental care was acknowledged.

In this trial, the selected interventions to assess their effectiveness reflect current practice. The strategies employed were passive dissemination strategies, audit and feedback (A&F) and computer aided learning with decision support (CAL-DS).

The passive dissemination strategies were the "mailing of guidelines and an opportunity to attend a postgraduate continuation education courses (PGCE)" that all groups were

exposed to. They were used as the control or non-intervention strategies. A&F and CAL-DS were selected as intervention strategies under investigation.

5.1.2a Guideline and Postgraduate Continuation Education Courses (PGCE)

In this study, all the participants received a copy of the SIGN guideline for the management of unerupted and impacted third molar teeth (SIGN 43, 2000) through the post. This is the most common method of disseminating guidelines in the United Kingdom. The SIGN guideline provided reliable information about the indications and contra-indications for the removal of third molar teeth. It was intended to serve as a tool by which scientifically valid and reliable standards of clinical management of unerupted and impacted third molar teeth could be implemented.

In addition, participating dentists were invited to attend PGCE courses to inform them of different clinical and radiological aspects of the assessment of third molar teeth based on the SIGN guideline (SIGN 43, 2000). The courses were available to all participating practitioners and their attendance was not obligatory, since this was a pragmatic trial. However, despite the provision of an incentive (accreditation for the postgraduate education allowance for section 63 courses *i.e.* NHS education for Scotland) and offering them the choice of three dates for the courses at three different regional centres, the attendance at these courses was disappointing (less than 50%). This seems to suggest that participants may have covered a number of other continuing postgraduate education courses related to third molar teeth such as those run by NHS Education for Scotland. So they may have thought that they would not gain any more knowledge on the same subject by their attendance on the courses. Or they were perhaps reluctant to

take part in additional activities especially if these were not compulsory. This is reflected in the findings by the Scottish Council for Postgraduate Medical and Dental Education Committee in 1990, that only 40% of respondents had attended in any Section 63 activity in the previous year while a further 40% had attended only one or two courses (Mercier *et al* 1998 *op. cit.*). However these findings are based on a sample of UK general practitioners over ten years ago and further research may well be necessary since the emphasis on the continuation of professional development as part of a Clinical Governance framework has been on the increase (GDC website: http://www.gdc-uk.org 2000).

All the groups including the non-intervention/control group were exposed to these passive dissemination strategies (*i.e.* mailing of the guideline and PGCE courses) to overcome the problems associated with groups that do not experience any intervention strategies (resentful demoralisation) (see Chapter 2, Section 2.3.4.1).

Passive dissemination of the guidelines and didactic courses are not effective in the uptake of scientifically valid recommendations and are unlikely to be effective in changing practice (Bero *et al* 1998). Nevertheless, utilisation of these two dissemination strategies reflects current normal practice in the UK and while their impact at best is small, they served to raise an awareness of the published guideline and were used as control interventions in this trial.

5.1.2b Audit and feedback

The Audit and feedback process emphasises the need for establishing best practice by attempting to improve patient outcomes by assessing practice against valid standards. Subsequently, the indicated changes are implemented to meet those standards and further review is used to validate improvement in the healthcare delivery. Despite the findings in medicine that the A&F strategy has a small to moderate effect (Effective Health Care 1999), it is commonly used as an intervention to improve the quality of care (MacKay and Thomson 1991, Mercer *et al* 1998) as funding for audit continues through purchasing NHS authorities. From April 2002, participation in dental audit became a requirement of the NHS terms and conditions of service for all GDPs. In Scotland, the Scottish Executive has made available a clinical audit allowance to support clinical audit within dental practices, while the SCPMDE (Scottish Council Postgraduate Medical and Dental Education) has the responsibility for promotion of dental audit within Scotland as part of a wider Clinical Governance framework (GDC website: http://www.gdc-uk.org 2000).

In this trial, audit and feedback (A&F) was selected as it was already established throughout the NHS for care providers to monitor their practice. The randomisation of participating practices across Scotland into different groups made it difficult for the particular groups intending to carry out an audit project. An attempt was made to make the task of meeting up easier for these participants by dividing them into audit subgroups according to the proximity of their practices to each other. Further, they were given a choice of carrying out the audit within their allocated groups or they could complete an audit on their own with support from the project team if it was not easy to

convene with the others in their group because of the distance. They were also given an opportunity to invite other interested practitioners in their area to take on the audit project together.

Previous experience and knowledge of the participants about audit was not clear so this was not assessed in this trial. But it has been shown in a study by Mercier and colleagues (1998) assessing the audit activity of 307 general dental practitioners in England that only 6% of the participants had an adequate knowledge of audit. This figure increased to 66% in a survey of 66 delegates (57 general practitioners and the remaining 9 working in other branches of the dental service) attending the Dentistry 2000 exhibition (Fyffe 2000). Hence to facilitate and co-ordinate the audit, one member in each group was selected as a facilitator by the SCPMDE Dental Tutor for Audit based on her knowledge of their previous experience and familiarity in carrying out an audit. This proved successful as the audit activities seemed well co-ordinated within those groups that managed to complete an audit project and no difficulties were experienced or reported.

The audit and feedback projects were conducted independently of each other on different aspects of clinical practice relating to third molar teeth and were supported by the researcher (MB) with help and advice from the Scottish Council Dental Audit Tutor.

Members of each group convened three times during their audit period. The facilitator gave feedback to the group by way of a summary of their audit results at the last

meeting. The researcher (MB) received a copy of the audit reports from the groups at the completion of their audit projects.

Of the 11 audit groups, seven completed their audit projects successfully while the other four failed to do so. Following consultation with the SCPMDE, Dental tutor for Audit it was considered to be the norm that not every practitioner chose to complete an audit project. This failure of some practitioners to complete their audit project could potentially or theoretically have had an adverse effect on the assessment of the effectiveness of the A&F intervention. However this is unlikely, since the majority of participants (21 (*i.e* 7 groups) out of 27 participants (*i.e.* 11 groups)) in the audit group managed to complete a project successfully. Nonetheless this trial attempted to represent a real life scenario and reflect daily realities.

No explanations were given by the practices when a reason for non-completion was requested. This may indicate that they were truly busy in their work or not motivated to participate or had an inadequate knowledge of the audit methods. This reluctance may however change as a result of the implementation of the Statutory Continuing Professional Development Programme announced by the General Dental Council (GDC website: http://www.gdc-uk.org 2000).

Retrospectively, it would have probably been more constructive if all the participants in the audit groups were invited to provide feedback and reflect on the outcome of their projects in a meeting. In this way, the quality of their feedback would be monitored and they could learn from each other's experiences and findings. On the other hand, there

was no guarantee that the practitioners in the audit groups, and especially those who did not carry out any such project, would have participated in these feedback sessions. The delivery of the audit and feedback intervention was exactly the same as to routine practice and reflects reality. Thus it is unlikely that failing to monitor the quality of a feedback session had any influence on the effectiveness of this intervention.

5.1.2c Computer Aided Learning with Decision Support (CAL-DS)

The decision support software was developed specifically for this trial with the aim of reinforcing the recommendations of the SIGN guideline (SIGN 43, 2000). It has the potential to assist dental practitioners in deciding on the appropriate treatment of third molars. The interactive nature of the CAL-DS programme enables the practitioners to explore the guideline at their own pace. The recent increase in the capacity of personal computers in dental practices and the rise in the use of the Internet was thought to facilitate the use of CAL-DS for the implementation of guidelines in the future. As shown by Fyffe (2000) in a questionnaire survey of 66 dentists, 81% of general practitioners use the internet to keep up-to-date with advances in dentistry (Fyffe 2000). In this study, in order to ensure standardization across different practices, 24 computers with relevant software for the CAL-DS programme were delivered to the participants of the study. An accurate assessment of the effect of the CAL-DS programme may have been complicated by the fact that some of the participants had limited computer skills. This factor has not been assessed in our study as this project was aimed to be a pragmatic trial and reflect reality. It is very unlikely that limited skills had any influence on the effectiveness of this intervention. since the use of software did not require much skill other than the practitioner being able to switch on the computer and use the

consumer friendly and self-explanatory software. Nonetheless the use of the laptop computer and software was demonstrated at the point of delivery (apart from one practice) and necessary instruction leaflets were supplied to all the participants in the CAL-DS group.

5.1.3 Methods of data collection

The study proceeded to collect retrospective cross-sectional data from clinical dental records before and after any intervention. Collecting data directly from the patients' dental records needed considerable logistical organisation such as negotiating access to all the participating practices in a geographically dispersed area and retrieving the case notes and extracting the relevant data.

However, the record keeping of participants was less complete than expected when it came to the assessment of a third molar tooth. Although the standard of record keeping of the participants was not assessed in this study, some of the case notes contained no dental charting, no periodontal diagnostic information and no radiographic report of third molar teeth. This resulted in further examinations of referral letters and other available sources such as existing radiographs on a frequent basis in order to extract possible relevant data missing from the dental practitioners assessment records and to prevent encountering any problems when it came to interpreting the data.

The evidence on current standards of record keeping in the UK tends to be anecdotal or circumstantial. However, the quality of the clinical dental records of 47 general practitioners who were voluntarily involved in a quality assurance programme by

BUPA Dental Cover in UK has been evaluated in a recent study by Morgan (2001). He reported that the quality of record keeping was poor and fundamental clinical entries that could impact on basic dental care provision were missing from many records. However these findings may not be a representative sample of the UK general dental practitioners (Morgan 2001).

Luckily, the poor record keeping of participating dentists in this trial had no impact on the interpretation of the data, apart from the longer time it took to search the other relevant sources (*i.e.* referral letters and radiographs) to collect the necessary information. Data collected such as the diagnosis of pathology, treatments prescribed and other clinical information were determined from the SIGN Guideline for the appropriate management of third molar teeth (SIGN 43, 2000). In addition, patients' demographic information and the reasons for their attendance were also recorded. The first set of data collected at the baseline provided unique information on the general practitioners' pre-interventions management of patients with third molar problems.

If the measuring instrument such as a data entry form or a questionnaire or the person collecting the data is altered from one data collection period to the other, it can introduce an "instrument bias" that may influence the outcome of the study and the effect observed may not be due to the intervention alone (Moles and dos Santos Silva 2000). This may be minimised or eliminated by the standardisation of the measurement process (Moles and dos Santos Silva 2000). In this study, a standardised data entry form was used to systematically record the data for both data collecting periods to prevent any "instrument bias". It contained short and closed-ended items, while remaining clear

and precise. This made the task of data collection relatively simple. In order to maintain a consistency in the information collected, the data entry form was not altered from pre- to post-intervention. This made the head to head comparison of collected data from both phases of the study easier and helped to eliminate any instrument bias. All records were retrieved and the data collected by researchers who were blind to the intervention groups to prevent any observer bias. All data (pre- and post-intervention) were collected by one researcher except data collected from 23 practices at the beginning of the pre-intervention data collection period. To prevent instrument bias, the second researcher was trained in the process of data collection by the first researcher.

5.1.4 Statistical issues

5.1.4a Sample size

One of the questions to be posed in the design of a study is the number of participants to include. It is an important question because if a study is too small it would not be able to answer a pre-determined hypothesis and would be a waste of time and resource and may give incorrect results by showing no effects when one exists (underpowered). However, a study should not be too large because resources would be wasted and ethical considerations would be raised as to whether fewer participants would have sufficed. Therefore, the main advantage of calculating sample size is to maximise the chance of detecting a statistically and clinically significant difference between interventions when a difference exists.

In most trials, the number of participants or the sample size is usually deducted prior to the start of the project. This can be estimated arbitrarily or can be calculated using

statistical methods. In this study, relevant advice about the sample size and the power of the study was sought from experienced statisticians and they calculated an a priori sample size. In an attempt to analyse the main effect of the interventions, it was estimated that a sample size of 60 practices collecting information on 240 patients was required to detect a 20% reduction in inappropriate extractions from 60% to 40% assuming 80% power and a 5% significant level (Casagrande *et al* 1978).

The projected estimate of inappropriate third molar extraction was originally set at 30% - 60% based on a study evaluating compliance in the UK with the USA's National Institutes of Health (NIH) criteria (Brickley *et al* 1996, II). It was estimated that 4-5 patients per practice would have extractions of their third molars during each data collection period based on data from the Scottish Dental Practice Board (Scottish Dental Practice Board 1997). This was believed to be an underestimation of the number of patients as it only considered surgical extractions in practice and not patients referred to specialised centres or teeth treated in ways other than extraction.

The unit of randomisation was dental practices rather than individual practitioners. This had implications for the sample size needed to achieve an acceptable statistical power (Kerry and Bland 1998). Since the outcomes for patients within any one cluster/practice were considered to be more similar than those across clusters/practices. To ensure that the study achieves the same power as it would have achieved without clustering, conventional sample size calculations needed to be inflated to correct for the intra-practice (intra-cluster) correlation (ICC) by incorporating an ICC into power calculation (Donner *et al* 1981). An estimate of the proportion of patients having third molar

extractions was considered for calculating the intra-practice correlation. However, there was a lack of available data for calculating the ICC for this condition and a number of assumptions have been made in the sample size calculation, assuming an estimate of probable intra-practice correlation coefficient of 0.1 which was drawn from comparable studies in medical primary care (Campbell *et al* 2001).

Nonetheless, in practice analysis of the study data has shown that the actual ICC in this trial was 0.15. This suggests that the study had a reasonably high power to detect a 20% reduction in percentages of non-compliance treatment if it existed.

The primary aim of the analysis was set to examine the main effects of the interventions, not their interaction effect. Using a larger sample size would possibly have provided greater power to allow for testing the interaction effect of the interventions. However, there was a trade-off between study power and sample size, since increasing the power would have resulted in an unrealistic sample size which might have threatened the feasibility of the study. This would have naturally required additional direct logistic and funding implications exceeding the financial limitations of the project.

5.1.4b The outcome measure

An outcome measure for this study was the proportion of patients whose treatment of their third molar teeth was in concordance with the guideline, *i.e.* treatments which constituted Evidence Based Practice (EBP). This was assessed by examining each case and comparing it with the SIGN guideline (SIGN 43, 2000). The treatments provided by

participating practitioners for patients with third molar problems at pre- and post-intervention were assessed independently by two clinical researchers in order to determine the degree of compliance with the recommended evidence (SIGN 43, 2000). Any disagreement between the two assessors was resolved by a discussion and an agreement was attained.

The reasons for having each case assessed by two researchers were to enhance the validity of the assessment and to reduce the possibility of subjective assessor-related bias while increasing the reliability of the outcome measures. Since human behaviour can be influenced by previous experiences, knowledge, expectations or beliefs (in research particularly) there is a chance that these biases influence findings. It is therefore imperative to have more than one assessor to prevent any subjectivity that might exist in assessment which may lead to a biased outcome (Day and Altman 2000).

Also, bias may be introduced if researchers assessing endpoints of interest are not blinded to the study status of participants. In this study, the two researchers who were involved in the evaluation of the outcome measure were oblivious of the type of intervention allocated to any particular practitioner in order to avoid detection bias.

The study produced evidence of good adherence to the guideline recommendations (EBP) at baseline (pre-intervention). The result also showed that aggregate EBP at pre-intervention was high (74%) with a statistically non-significant improvement of 4% on this level of compliance (78%) observed after the interventions. The guideline adherence (EBP) for different intervention groups showed that EBP increased in all

groups following interventions; nevertheless these increases did not seem to be statistically significant (see Section 4.1.2, Table 4.1.10).

In the first instance, this result appears to suggest that the baseline level of compliance is considerably higher than those reported at the time of planning the study by other researchers (Brickley *et al* 1996, II) which sparked many of the guidelines' concerns. Inappropriate third molar extraction in the UK was assessed by evaluating the degree of the treatment's compliance with the USA's National Institutes of Health (NIH) criteria (Brickley *et al* 1996, II). The a priori estimate of 40% adherence to the guideline at baseline for this trial was based on their estimate of the numbers of extracted third molar teeth (Brickley *et al* 1996, II) as opposed to the overall management of third molar pathology. Conversely, the main outcome measure of interest in the current trial was dentists' adherence to the guideline recommendations (appropriate management of third molars) and the adherence of the sample population to guidelines proved to be substantially higher than expected. The possible explanation may be that the high adherence of dentist to guideline recommendations could be because the dentists were more confident in the management of other third molar teeth predicaments such as caries, periodontal disease and pulpal pathology and the main predictor in compliance with guideline was the type of cases that are presented to the dentists. It may be possible that the findings from the above publication (Brickley *et al* 1996, II) and others have influenced knowledge levels and practice. There have been discussions on the appropriateness of the extraction of impacted third molars in various dental journals over recent years. One of the examples is the Effectiveness Matters publication from the Centre for Reviews and Dissemination at the University of York which was published

in 1998 (Effectiveness Matters 1998). The findings of the review were based on other reviews and research findings that have been widely reported in the dental press (Lands 1998, Song *et al* 1997, Daley 1996, Mercier & Precious 1992). Also there were a number of different third molar guidelines such as the Royal College of Surgeons of England Faculty of Dental Surgery (1997) and NICE (1999) which were published in recent years. These may have influenced Scottish dentists' behaviour. Data from NHS Scotland's Management Information and Dental Accounting System (MIDAS) showed a reduction in both surgical and non-surgical wisdom tooth extractions took place in Scottish General Dental Practices before and during the experimental period (Graphs 4.5.1 and 4.5.2). However, each randomised group experienced the same global pattern of history and the rate of extraction and the level of compliance of practitioners did not differ significantly between the randomised groups. It appears that randomisation was successful and there is no reason to believe that history bias had influenced the practitioners' behaviour in each intervention group. Nevertheless if publications relating to the third molar teeth have influenced the care of patients, it is one of the few examples of relatively passive dissemination and implementation being effective in altering clinical practice. It is likely that the high level of compliance at baseline has produced a "ceiling effect" (Grimshaw *et al* 2000) which is that the high performance scores may indicate that there is little room for improvement. The self-selected group of volunteer dental practitioners who consented to participate in our trial may not be representative of the population sample as shown by comparing the third molar extraction rates of the sample population of dentists with the other dentists in Scotland in Graph 4.5.5 using data from MIDAS. This could have then resulted in an overestimation of the population level of compliance at pre-intervention. Another

explanation for the high compliance could be the "Hawthorne Effect" (Randall & Cebul 1991, Grimshaw *et al* 2000). Since the practitioners are aware of being assessed they might therefore perform more effectively than at other times.

5.1.4c Data analysis

The aim of the primary analysis of data was to examine the main effects of the interventions. In this study, data were analysed using two general approaches for analysing data from a cluster randomised controlled trial, *i.e.* a cluster level analysis and patient level (multilevel) analysis. As has been already mentioned in the method section, an analysis of the covariance model correcting for baseline compliance was originally considered but there was a weak correlation between pre- and post- cluster level compliance rates. In addition, visual inspection of the weighted pre-intervention mean of percentage compliance for different groups did not highlight any notable baseline imbalance (see Chapter 4, Table 4.11). Therefore, the main analyses were performed on the post-intervention data only.

Data were analysed at the cluster level using a 2x2 factorial design in order to study the main effects of the interventions per protocol. A summary measure for each cluster was calculated. As a result, each cluster was represented by one single data point, *i.e.* cluster weighted mean, using simple statistical tests for significance such as the t-test. The level of EBP between the groups was analysed using a weighted t-test. Since the estimates from dental practices with larger numbers of patients were more precise, the practice EBP estimates were weighted by the number of patients seen in that practice (Bland and Kerry 1998). The cluster level approach assumes that the data from each cluster is

independent and the intra-cluster correlation (ICC) is effectively equal to one. However, this is rarely the case as shown in a study of UK data sets relevant to implementation research (Campbell *et al* 2001). They showed that in a primary care setting, the ICCs for process variables, such as adherence to guidelines or compliance with best practice recommendations, appear to be of an order of 0.05 to 0.15. When the intra-class correlation is less than one but greater than zero, the effective sample size is less than the total number of subjects in clusters but greater than the total number of clusters. Thus, analysing by the unit of randomisation at cluster level analysis, while having the benefits of simplicity, reduces the power of the study to detect significant effects and does not make the most efficient use of the available data (Campbell *et al* 2000). Therefore, analysing as two separate trials as factorial design implied (1. CAL-DS versus no CAL-DS and 2. A&F versus no A&F) assists in making the most of the available data.

At the patient level approach, EBP was analysed using multilevel modelling (see Chapter 3, Section 3.6.2.3b). The patient level analysis allows adjustment for the hierarchical nature of the data (patients nested within dental practices). The advantage of this model is that the effect of confounding factors such as heterogeneity in patients (case-mix) was accounted for in the analysis since controlling for "case-mix" or "confounding factors" such as the different type of patients' characteristics will affect the outcome. (For example: suppose a dentist is always more likely to follow the guidelines if the patient has caries. Different rates of caries treatment across dentists will lead to different rates of compliance with guidelines in a manner that is completely independent of the effectiveness of the interventions.) In addition, this method utilises

all the patient data, while modelled for the inherent correlation within clusters, thus increasing the statistical power of the analysis (Campbell *et al* 2000).

Originally, the multilevel model in the study used all data with a variable to indicate whether a case was pre- or post-intervention and the result of comparison (pre- to post-) was non-significant (see Chapter 3, Section 3.6.2.3b). This was an additional reason for focusing on post-intervention data in the main analysis. All possible variables and interactions between time, intervention and patients' characteristics or case-mix were modelled. But, the cells in the regression model became very small and many of them were empty due to the low number of cases. This can result in unreliable standard errors. For example, there were no non-compliant caries cases and modelling these cases gave rise to a situation analogous to an empty cell in a 2X2 contingency table. Therefore, when the model was reconsidered it was considered necessary to discard all cases which were low in number.

Nonetheless, both approaches allowed the effect of the interventions to be tested using post-intervention data and gave very similar results. Although the hierarchical modelling handled an extra source of variation it did not explain any other differences between the groups.

To assess whether the interventions had a different impact within any group, a secondary sub-analysis at the patient level considered the variance of EBP based on both pre- and post-intervention data. Four separate groups (i.e. non-intervention, Audit, CAL-DS, Audit and CAL) were modelled for this sub-analysis and the trial was powered on a factorial design. Therefore, the results of this approach should be

considered with caution. Also the problem is with the low number of the cases in regression model. Standard errors can then become unreliable and the confidence intervals can become too narrow. In sub-analysis, when the model was considered, none of the cases which were low in number were discarded.

The analysis was performed on the intention to treat (ITT) basis considering the pragmatic nature of the study. All dental practices were analysed according to the groups into which practitioners were randomised, whether or not the participants utilised the guideline implementation strategies fully (Newell 1992, Prescott 1999). For example, many of the practitioners invited to participate in a PGCE course did not attend or practitioners in the A&F group did not carry out an audit project. Therefore all the dental practitioners participating in the intervention phase were analysed in the group to which they were initially randomised regardless of whether or not they utilised the intervention (Roland and Torgerson 1998). So, if the analysis had been conducted by excluding dental practitioners who did not utilise the interventions that they were randomised to, there would have been a possibility of overestimating the effects of the interventions in real world settings and introducing bias in the conclusion (Newell 1992).

5.1.5 Validity

The potential influence of each source of threat to the validity of this study is considered and the two main types of validity in relation to the design of this study are assessed.

5.1.5a Potential threats to internal validity

Appraising the findings of a study necessitates the evaluation of its internal validity. This enables the researcher to assess the degree of confidence that one can have in its results. Cook and Campbell (1979) provided a list of common threats to the validity of a study (see Chapter 2, Section 2.3.4.1-2) and the likely impact of these threats on this trial has been considered.

The design (*i.e.* RCT) adopted in this study protects against many of the internal validity threats. When participants were randomly allocated to each intervention group, the composition of each group on average was similar (Section 4.1.2, Table 4.1.10). Therefore, there should be no selection, maturation or selection-maturation biases (see Chapter 2, Section 2.3.4.1).

Participants experienced the same testing conditions and research instruments (such as assessor or observer, data collecting forms). As a result there should be no testing or instrumentation biases. Each group experienced the same widespread pattern of history so no problem of history bias should exist either.

However, the randomisation does not take care of threats such as the imitation of treatment, resentful demoralisation in groups receiving less desirable treatments and compensatory rivalry. In this study, an attempt was made to avoid contamination and imitation by randomising at practice level to ensure that practitioners in one practice would receive the same interventions. However, possible contamination outside the practices through informal socialising, if it existed, could not be prevented or accounted

for, although there was no evidence to suggest that these occurred since there were no sudden changes in practice observed between the two phases of the study.

All the groups including the non-intervention/control group received a copy of the guideline and a chance to attend a PGCE course to minimise the adverse effects of resentful demoralisation. In this study, as with any other investigation, it is hard to assess whether compensatory rivalry has taken place or not.

Further, the validity of the statistical conclusion was assessed as an integral part of addressing the issue of internal validity. In this study, attempts have been made to overcome most of the threats to internal validity. An *a priori* sample size calculation was carried out prior to the recruitment stage which was achieved in order to assess the main effect of the interventions. An *a priori* analysis strategy was considered in the protocol to reduce the possibility of "fishing". A multilevel modelling was used to correct for the effects of variables associated with the outcome (see Chapter 3, Section 3.6.2.3b for further details).

A further secondary sub-analysis using general linear modelling was carried out at a later stage of the trial which could possibly give rise to "fishing" or Type I error. When multiple statistical tests are performed on the same data it can be expected, by chance alone, that some will show an apparently enhanced effect. This is one of the reasons that the results of secondary sub-analysis should be considered with prudence.

One way of controlling for unreliability of measures is to use aggregated units, e.g. groups instead of individuals (Cook and Campbell 1979), as was employed in this study by randomising the participants into groups. All participants in the study utilised the interventions uniformly as proposed by the study design, apart from some of the dentists in audit groups. The failure of the dentists to carry out an audit project reflects reality and the pragmatic nature of the trial.

The random irrelevancies/confounding factors in the experimental setting such as contamination between experimental groups were avoided by randomising at practice level rather than dentist level which might influence the endpoints of interest. To guard against any violation of assumptions of statistical tests, the groups to be compared were required to be equivalent in baseline variable such as characteristics of the patients and treatment's compliance with guideline which was checked by comparing pre-intervention groups. An attempt was made to reduce the heterogeneity of subjects across experimental groups by correcting all the possible confounding attributes of the regression model.

5.1.5b Potential threats to external validity

External validity of a study is imperative in the generalisation of its findings to a given target population (see Chapter 2, Section 2.3.4.2). For assessing external validity of this study, related factors such as target population and interventions were considered.

5.1.5bI Target population

The study sample should be representative of the population sample to which we hope to generalise the findings, and to draw up such a sample according to Cook and Campbell (1979), one should follow a two-stage randomisation process. In the first stage (which is called random sampling for representativeness model) samples are randomly chosen from the population and subsequently these randomly selected units are randomly allocated to various intervention groups. However, on the whole this is a sample of convenience, i.e. volunteers who most often give consent for continuing with the research study even in this model (Cook and Campbell 1979).

In this study the participating dental practitioners were volunteers who consented to participate in our trial and possibly had a particular interest in research or the subject of the research. They were perhaps more likely to participate in research than others. It may be plausible to assume that subjects who declined to participate in the trial may have different characteristics to those who did participate, rendering the sample less representative of the target population and possibly introducing bias (Cook and Campbell 1979).

Nevertheless, this is a pragmatic trial in a primary care setting in which a subpopulation of participants were successfully randomised. The analyses of the data are reasonably high in internal validity but perhaps low in external validity. The reason behind the low external validity is informed consent. According to the recommendations of ethical guidelines for the conduct of randomised trials, each individual subject in a study should give an informed consent to participate (Edwards *et al* 1998) and they should be provided with enough information about the study to allow for a well-informed consent. Accordingly, there was a systematic tendency in this study to recruit only those subjects who consented and volunteered to participate in the trial. This is one of the acute problems of research in primary care settings.

Data from other surveys provided the opportunity to compare the characteristics of our dental practitioners with those who participated in other studies. The characteristics of our participating dentists display a similarity to those in two UK surveys. One was based on a self-completion questionnaire carried out by the British Dental Association (BDA) in the UK (BDA Dental Business Trends 2000) where the study sample was a population of dental practitioners who were members of BDA. In this survey, more male dentists (68%) participated than females (32%). Sixty-eight percent of their dentists qualified in the last 23 years. The other survey was also based on a self-completion questionnaire carried out in Scotland by the Department of Public Health, University of Aberdeen and a higher percentage of male dentists (66%) responded to their questionnaires than females (34%), (Russell *et al* 2000). The participants in both of these surveys were volunteers. In this study, a greater number of volunteer dentists were male (80%). Fifty percent of participants qualified in the last 16 to 24 years.

Cook and Campbell (1979) hold the opinion that even if "we leave aside the important ethical problems and look only at pragmatic issues, it may be self-defeating to conduct one's randomised research today in a way that directly or indirectly demeans the respondent. Of course restricting the population to volunteers would decrease the external validity but it would not restrict personal freedom. The internal validity could be enhanced by carrying out a randomised controlled trial over other design alternatives for honestly assessing the relative effectiveness of various interventions".

In this study, there was an inevitable trade-off between one type of validity and the other and internal validity was best served by carrying out a randomised controlled trial; however, the participants prepared to consent to take part were probably less representative of the population sample. The priority ordering for many studies including this trial that is concerned with testing "whether the interventions as implemented caused the desired effect" is internal validity, external validity and statistical conclusion validity (Cook and Campbell 1979). This trial seems to score high in 2 out of the 3 validities that is high in internal validity and statistical validity.

5.1.5bII Implementation of interventions

One of the reasons for the problem of heterogeneity in the way that interventions are implemented is the personal effort the intervention requires (Cook and Campbell 1979). In this study, the problem of intervention heterogeneity was minimised but not eliminated by standardising the delivery of the CAL-DS programme by computer since there might have been some variability in the frequency of exposure to the intervention because of absence or differences in the motivation to use the computer software by participants. Interventions are more likely to be standardised if they are delivered by

trained researcher personnel (Cook and Campbell 1979). However, it would require additional logistical requirements and funding requirement and would not reflect a real life situation.

5.2 Interpretation of results

5.2.1 Recruitment

The recruitment of a sufficient number of participants to a study is an essential measure in order to have enough power for statistical analysis. This study, like any other in primary care, had acute problems when it came to recruiting a sufficient number of practitioners to be a sample representative of the population and also to maintain the motivation of all the recruited practitioners to complete the study. Charlson and his colleague have shown that 66% of trials never achieve their projected sample size (Charlson and Horwitz 1984). In this study, the projected sample size was 60 practices (minimum number of 240-300 patients). The *a priori* sample size was achieved at the beginning, since 63 of the 565 (11%) invited dental practices volunteered to participate in the study. As some loss of subjects in any study is to be expected, the three additional practices consented above the 60 estimated to be required by the power calculation was thought to be sufficient. However, 12 of these withdrew before pre-intervention data collection and randomisation. An additional four practices withdrew between the two data collection periods, leaving 47 (8%) practices. Several reasons were given for the withdrawals. They were too busy to participate, not interested, changing practice, refurbishing their practice or simply changed their mind. The reasons given for withdrawal from the study indicate that they were principally for reasons other than the demands of the project. Comparison of the characteristics of the participants (such as

age, gender, postgraduate qualification and allocated intervention group) between those who had withdrawn from the study and those who had continued showed no significant statistical differences. These characteristics of the participants did not seem to have any influence on their continuation with the project.

Luckily, the analysis of the main effect of the intervention was not compromised by the number of practices that withdrew since in practice it was shown that the study had a reasonably high power to detect a 20% change in the compliance of participants with the guideline if it existed (see Section 5.1.4a).

In this trial, an attempt was made to encourage recruitment and to avoid withdrawal once practitioners had consented to participate. Researchers were employed for collecting the data to minimise the disruption to daily clinical activities and as an incentive and a sign of appreciation the practices were given a voucher at the end of the data collection at each pre- and post-intervention session.

The participant practitioners were requested to register the name of all patients in the 16-24 age bracket with any dental problems, in order to pre-empt introducing any biased registration or biased treatment. Three thousand, three hundred and forty-two patients were registered at the pre-intervention stage as compared with 1935 patients recorded at the post-intervention stage. However, the number of patients with third molar problems at post-intervention was twice the respective number at pre-intervention (426 versus 244). An explanation was not found for this by the author. Nevertheless if this is the case it is unlikely to affect the findings, as the study examined how these

patients were cared for and not the identification of the patients. However, these figures may be speculated on, possibly one of the reasons that the practitioners became more acutely aware of the aim of study following the pre-intervention data gathering and intervention phase and this resulted in a higher number of patients who were registered with third molar problems at the post-intervention (testing bias). The randomisation should have eliminated this bias as well as the attempt made to prevent this threat (testing bias) by keeping the process of data collection as unobtrusive as possible to minimise many reactive-based threats to the internal validity. As mentioned above, this was done by asking the practitioners to record the name of all the patients who came to see them with any dental problem. Or this finding could be that the dentists participating in the trial were more confident in their management of the condition following the interventions, if not there was an improvement in case-note keeping as a consequence of their participation in the trial.

In order to search for more explanations, the proportion of patients who had a third molar tooth removed by Scottish General Dental Practitioners over a 36-month period (January 1998 to January 2001) were compiled from MIDAS (Graph 4.5.3, 4.5.4). We were hoping this would enable us to detect any seasonal changes in the proportion of attendees to these practices over the data collection periods. However, no discernible reduction in the proportion of these patients was detected after comparing the two data collection periods (Graph 4.5.3, 4.5.4) so with that result we could not explain the large variation observed in the study.

5.2.2 Intervention effectiveness results

The result of the main research question "the effectiveness of the main interventions (A&F and CAL-DS) under investigation in altering evidence based practice (EBP) for the management of third molar pathology" will be discussed in this section. The outcome of the assessment of the post-intervention data by both the cluster level analysis and multilevel analysis are also discussed. In addition the findings of the sub-analysis, which examined both pre- and post-intervention data to assess the relative changes in EBP within the groups following the interventions, will be considered. Finally the results of analysis of the patients' characteristics and the SIGN guideline recommendations about recording medical history and radiographic examination in this study are discussed.

The intervention groups were examined for any imbalances at baseline. The result demonstrated no evidence of any notable baseline imbalances and there were no significant differences between the groups at pre-intervention whatever type of analysis was used, when no control for the case mix in the model was considered. It appears that the randomisation process has balanced the confounding factors across the groups. Conversely, in sub-analysis, when consideration was given to the case-mix in a study model to examine the effect of different intervention groups (i.e. non-intervention, A&F, CAL-DS, A&F + CAL-DS) it became apparent that the groups differed at the baseline. Specifically, CAL-DS was different from other groups and significantly different from the A&F group.

This baseline imbalance could have a significant implication for the sub-analysis results, since imbalance between groups could bias the outcome and the statistical tests (chance bias). It (see Section 2.3.2) seems to suggest that the baseline groups may not be comparable and may not experience the same secular trends or sudden changes. Hence, the observed effect of the CAL-DS intervention could by chance be due to baseline variables such as different characteristics (i.e. caries, pericoronitis) of the patients seen (Grimshaw *et al* 2000). However, to avoid this bias, allowance was made for baseline variables in the statistical analysis. Therefore it is likely the chance bias has been avoided.

On examination of the post-intervention data for the main effect of the interventions, both the cluster level analysis and the hierarchical (multilevel) analysis gave very similar results. The main finding of the trial was that there was no evidence of a significant effect for either A&F or CAL-DS ("A&F versus no A&F" and "CAL-DS versus no CAL-DS") implying that neither of these interventions was more effective than the mailing of the guideline and attendance at the postgraduate course.

In sub-analysis, examination of the relative changes in EBP within the groups (i.e. non-intervention/control, A&F, CAL-DS, A&F + CAL-DS) from pre- to post-intervention was made. There again, the results showed non-significant changes in EBP for all intervention groups, implying that no evidence of any statistically significant effect of the interventions within their own group was found. However, when controls for patients' characteristics were included a change from baseline was significant for the CAL-DS group at post-intervention and that the CAL-DS intervention seemed to have

worked in this group, resulting in an increase in compliance with the guideline (EBP). The likely explanation for this finding is that the intervention groups only differed on case-mix at baseline. At pre-intervention the dentists in CAL-DS group treated only 9 (18%) cases of caries. At post-intervention, this number rose to 22 (25%) cases (see Chapter 4, Section 4.4.4). This proportional rise of cases of caries from pre- to post-intervention seems to be the reason behind an apparent effect of CAL-DS intervention observed in this group. The treatment of caries in both periods was highly compliant with the guidelines among all dentists (see Chapter 4, Table 4.4.27). Therefore, the effect of case-mix (i.e. higher proportion of caries cases) resulted in the higher compliance at post-intervention and not the effect of CAL-DS intervention per se. In addition, as explained in the statistical validity section, when multiple statistical tests are performed on the same data it can be expected, by chance alone, that some will show apparent enhanced intervention effect (fishing or type I error).

In sub-analysis, the analysis of the effects of the interventions demonstrated that the interventions (i.e. A&F, CAL-DS, A&F + CAL-DS) would have no more impact than the current practice of mailing guidelines and postgraduate courses (non-intervention group/control) in altering evidence-based practice for third molar management. No significant statistical differences were detected when the changes "the pre- and post-intervention EBP differences" between the interventions groups (when controls for patients' characteristics are included or excluded) were compared, implying that interventions under investigation had no statistically significant effects in this study and certainly had no more impact than the current practice of mailing guidelines and offering postgraduate courses.

In the analysis (sub-analysis) of the effects of the current practice of mailing guidelines or postgraduate courses, it is known that the current practice does not significantly influence the compliance of the recruited general dental practitioners (Bero *et al* 1998, Davis *et al* 1999). Similarly, the study found a non-significant effect of these current interventions. It is clear that passive dissemination of guideline and didactic educational courses, in which information about the SIGN guideline was presented to a passive audience, were unsuccessful. These strategies may have raised awareness of a published guideline in this study but they were generally ineffective in altering the practice, no matter how important the issue or how valid the methods were.

In the main analysis, the interaction between interventions (i.e. the combined effect of CAL-DS and A&F) was modelled and it too was not statistically significant. Similarly, in a sub-analysis examining the change within CAL-DS+A&F group no effect was seen of this combined intervention. These observations may not be a true representation of the interaction effect since the study was not set up to have enough power to detect any interaction effect of these interventions if it existed. The analysis was carried out to satisfy the statisticians' inquisitiveness in case the analysis was different. It is a fact that further analysis could give rise to the possibility of the "fishing" error. But human beings cannot be stopped from being speculative.

The main finding of this trial "that neither A&F nor CAL-DS was effective in increasing the general dental practitioners compliance with the guideline" is not unique to this study. These interventions (i.e. CAL-DS and A&F) have been evaluated in medicine before. A trial recently published by Eccles and his colleagues (2001)

suggested that A&F does not change behaviour for primary-care radiology referrals (Eccles *et al* 2001). However, a small to moderate effect of A&F was reported in the systematic review by Thomson O'Brien and his colleagues (2002), although they had reservations about the inappropriate level of randomisation or analysis of the studies that they examined, as these concerns could cause an overestimation of effect size (Thomson O'Brien *et al* 2002, I &II). Nevertheless A&F is a potentially fundamental strategy as already most care providers in the NHS are familiar with the process of the audit, especially now that audit is mandatory as a part of the NHS terms and conditions of service and funds have been made available to support clinical audit within dental practices. It is a potentially valuable intervention as a starting point, which could provide baseline information from which more effective implementation strategies can be employed.

An evaluation of CAL to develop clinical decision-making skills in primary dental settings showed that it had no effect on dentists' treatment decision-making behaviour (Kay *et al* 2001). A systematic review of 98 studies on computerised information systems found that different interventions, including provider and patient prompts, computer-assisted patient education and computer-assisted treatment planners improved patient care (Balas *et al* 1996, I). Other reviews of relevant trials showed that such systems could improve the performance of practitioners in terms of decisions on drug dosage, the provision of preventive care and clinical management of patients but not in diagnosis. More work is needed on patient outcomes (Johnston *et al* 1994). Similarly, a review of 68 studies showed that the use of computer-based decision support systems

has also led to improvements in decisions on drug dosage, provision of preventive care, practitioner performance and patient outcomes but not in diagnosis (Hunt *et al* 1998).

Nevertheless, CAL-DS may have potential as an educational tool and as a reminder and decision support system, since the practice of dentistry carries an obligation to life-long learning and dentists are facing information overload. They often need to access information which is timely and specific to help them decide on the appropriate management of their patients. It is often impractical to access printed information at the time of consultation to help their decision-making. In addition, the NHS invests lots of resources into continuing professional development. Within most dental practices, dentists are increasingly using computerised technology in day-to-day clinical practice and a computerised decision support system would be easy to implement. These could have important implications for promoting the use of a CAL-DS programme by the profession.

5.2.3 Results of descriptive analyses of patients' characteristics

Brickley and his colleagues suggested that the most frequent disease related to retained lower third molar teeth is pericoronitis (Brickley *et al* 1995) which generally affects young adults aged 16-25 with the incidence of this disease decreasing very sharply beyond the age of 30 years (Ahlqwist and Grondahl 1991). The most frequent pathologies diagnosed for patients in the 16-24 age bracket were pericoronitis and caries in both phases of the study.

Pre- and post-intervention comparison of treatments provided for patients with a third molar problem over all groups revealed that more third molar teeth were restored or kept under observation after intervention (Section 4.4.1.2, Table 4.4.12). For those patients who had an extraction, a statistically significant reduction was detected between the proportion of third molar tooth extractions at the pre-intervention (37%) compared to the post-intervention period (27%) over all groups and the rate did not differ significantly between the groups. This overall effect cannot be attributed to the effect of the interventions, as this was already confirmed by statistical non-significant change of the overall compliance (EBP) of dentists with the guideline (SIGN 43, 2000) (see Chapter 4, Table 4.1.9). However, it is possible that the participation in the trial and awareness of the guideline recommendations might have had an overall effect in reducing the number of extractions independent of different interventions and possibly the participating practitioners became more confident in their management of the condition. This overall change in the rate of extraction could be perceived as an effect of history bias, since a concomitant reduction in third molar tooth extractions occurred in Scottish General Dental Practices during the experimental period (history) (Graphs

4.5.1, 4.5.2) as has already been discussed (see Section 4.1.4b). The randomisation appears to have been successful in eliminating this bias since there was no statistically significant difference in the rate of extractions of the third molars between the groups. Another plausible explanation is that the change observed in the overall proportion of third molar extractions is nothing but the differences between the relative frequencies of certain pathologies requiring the treatment of extraction (i.e. case-mix such as pericoronitis and caries) between two periods of the study. Pre-intervention, 62% of patients had pericoronitis and 30% had caries. Post-intervention, there was a statistically significant drop in the percentages of pericoronitis cases (54%) and a non-significant increase of caries cases (32%) (see Chapter 4, Table 4.4.1).

Equally, examining the individual third molar pathology and assessing the treatments of patients with pericoronitis revealed that post-intervention observed and reviewed cases became more frequent relative to referral and extraction cases. The results seem to suggest that following intervention, dentists carried out fewer extractions or referrals for patients presented with pericoronitis and kept patients under regular surveillance more often. A large number of treatments provided for pericoronitis followed EBP whether they were observed at pre- or post-intervention. Treatment of pericoronitis did increase in compliance but this was not statistically significant (see Table 4.4.25) and the effect was not different across the groups (see Section 4.2.1).

The result shows that dentists did not have any problem in interpreting the guideline recommendation about the status of infection, where the guideline strongly recommends the removal of any symptomatic wisdom tooth especially where there have been one or

more episodes of infection such as pericoronitis. The author thought this recommendation could be clearer if a specific explanation about the severity of infection with one single episode of presentation was added, since the optimum management of third molars which have been associated with only one single episode of pericoronitis will depend on the severity of symptoms. Less severe infection and those patients that present with a history of a single episode of infection which resolves spontaneously may be better managed with antiseptic therapy and regular surveillance rather than removal. The results suggest that the authors concern was not shared with the participating dentists, as they carried out fewer extractions for patients presented with pericoronitis and kept them under regular observation more often. Especially, there was no statistically significant change in compliance of the treatment (EBP) of patients who had recurrent pericoronitis in comparison with the single episode (P=0.4) (see Chapter 4, Section 4.4.1.1.a). The result suggested that the recurrent status of pericoronitis was not a strong indicator of EBP-treatment in this study (Table 4.4.7, 4.4.8).

Investigating the treatments for patients with third molar caries shows that treatment of caries with restoration became more frequent relative to extraction after intervention. The results seem to suggest that following intervention dentists carried out a higher number of restorations and less extractions for patients who had caries. All treatments provided for caries appear to have followed EBP whether they were observed at pre- or post-intervention stages.

The same explanation (apart from case-mix) as is given above for differences observed in the proportion of third molar extraction, applies here for cases with pericoronitis or caries. It is obvious that the differences observed are not explicable by case mix since we considered each pathology variable individually.

5.2.4 Association of gender with tooth pathology

The sex specific tooth pathology in this study is similar to those reported previously. The Adult Dental Health survey in the UK in 1998 demonstrated that men on average had more decayed teeth than women. They made a similar observation for the 16-24 age range, which is identical to the age range of our study population (Adult Dental Health Survey 1998).

Brickley and Shepherd found that females were more likely to have third molar problems than males (Brickley and Shepherd 1996). Similarly, the data for this study showed that females were more likely to have pericoronitis than males. On the other hand, males are more likely to have caries in comparison with females. These findings may be accounted for by the fact that more women regularly visit a dentist than men and would thus be more likely to be diagnosed with a pathology. It is possible that since men have a higher number of lower first and second molar teeth removed (represented by higher DMF scores) that they have a lower rate of third molar teeth impaction and do not frequently have problems with their third molar teeth (Brickley and Sheperd 1996).

5.2.5 Results of the analysis of recording of a medical history and a radiographic examination while taking into account the SIGN guideline recommendations

In the next two sections, additional findings from the data are discussed even though they are not part of our main study question but could be important since they are relevant to recommendations in the SIGN guideline (SIGN 43, 2000).

5.2.5.1 Availability of a radiograph and agreement with guidelines

The SIGN guideline (SIGN 43, 2000) recommend radiographic examinations to provide the information necessary for adequate assessment of third molar teeth, especially prior to their removal. However, a radiographic examination is not required at the initial examination and routine regular radiographic examination of third molars is not recommended. Hence, as part of SIGN guideline recommendation the availability of radiographs and the rate of adherence of treatments to the guidelines was examined.

The results seem to suggest that when participating dentists carried out a radiographic examination, a high proportion of the patient treatments followed EBP in both pre- and post-intervention. Although the proportion of radiographs available was relatively small (12%), it is likely that radiograph examination did help the dentists with decision making about the type of treatments. The low percentage of radiographs could be due to the recommendations that a radiographic examination is not required to be carried out at initial examinations and routine regular radiographic examination of third molars is not recommended.

5.2.5.2 Medical History

The SIGN guideline (SIGN 43, 2000) recommends that a full medical history is recorded during the initial assessment of patients with third molar problems since medical history is imperative for decision making about the removal of third molar teeth. Certain medical conditions can render the removal of third molars an unacceptable risk to the overall health of the patient (e.g. leukaemia) or in direct contrast with the medical scenario when removal is advisable as the risk of retention of the tooth outweighs the potential complications associated with extraction (e.g. prior to radiotherapy or cardiac surgery).

When the recorded medical history was examined with the comparison of pre- and post-intervention, it was revealed that a higher proportion of patients with a recorded medical history were found in post-intervention than pre-intervention. It is likely that although the interventions did not have any effect in changing the compliance of the sample population with the guideline, they possibly had a positive overall effect (not individual effect) on improving the record of medical history with practitioners becoming more organised and keeping better records at post-intervention.

5.3 Summary and Conclusion

This study has been one of the first rigorous evaluations of intervention strategies in the primary dental care setting which adopted the more demanding cluster randomised controlled study design and attempts have been made to be rigorous in the quality of conduct. Examining the quality of current RCTs in dentistry, a recent systematic review revealed that the quality of RCTs in periodontology, judged by their publications.

frequently did not meet recommended standards (Montenegro *et al* 2002). Also, a systematic review of the quality of 43 RCTs evaluating the effectiveness of oral implants found the reported RCTs to be of a poor quality (Esposito *et al* 2001). Both of these reviews are related to the quality of reporting and not the actual conduct of the trials. Nevertheless, if the quality of reporting does reflect the actual study conduct, there could be a significant impact on the outcomes of these trials (Richards 2003).

This trial provides the best estimate of effectiveness of the interventions currently available in dentistry. This rigorous study examined the main effect of interventions and specifically showed that there was no statistically significant effect of either CAL-DS or A&F on the implementation of a SIGN guideline (43, 2000). Nonetheless, the study produced evidence of good adherence with the guideline recommendations (EBP) at baseline which is encouraging. The sample population's guideline compliance proved to be substantially higher than expected. There are several possible reasons for this finding. The high level of compliance could be due to the influence of publications and guidelines related to the subject of third molar over the last 7 years that has raised the awareness of all dentists to the management of third molar condition over recent years (NIH consensus criteria 1980, The Faculty of Dental surgery of Royal College of Surgeons of England (1997), Effectiveness Matters 1998, NICE 1999). Or it could be that very few guidelines have been published in dentistry in the last decade. This low number of published guidelines possibly helped to focus the dentists' attention. For instance, so far the SIGN has published only 2 guidelines in dentistry as opposed to over 40 guidelines in medicine. The dentists were possibly more aware of the few new published guidelines in their field and this may have influenced the high guideline

adherence. It is even possible that the participating Scottish dentist who used the guideline may well have had a sense of ownership since the SIGN guideline (43, 2000) was published in Scotland. This may well have resulted in a greater extent of implementation of these guidelines as familiarity with the guideline developers may influence the perceived appropriateness of the guideline and compliance of the participants (Feder *et al* 1999).

The participation of the volunteer dentists who consented to take part in the trial may be a significant confounder explaining the high compliance. These volunteers may manage their patients more effectively rendering the sample more compliant than the target population. Other explanations for the high compliance could be the "Hawthorne Effect". Practitioners may have acted differently because they were participating in the trial. An attempt was made to reduce the Hawthorne effect though, by keeping the process of data collection as unobtrusive as possible.

Therefore, the lack of difference between the interventions may be a true lack of effect or it may be a reflection of the high level of compliance at the baseline which may have produced a "ceiling effect", where no greater improvements within the group were possible.

Demand for high quality research in primary dental care is growing, especially for evidence from randomised controlled trials. These studies are complex and expensive to conduct. Failure to recruit participants from a full range of the general dental practice population may reduce the generalisability of such trials' conclusions. The inevitable

bias inherent in this study in collecting data from volunteer dental practitioners (selection bias) may restrict generalising its relevance in other settings.

However, it is an unrealistic expectation to assume that this pragmatic study can have all the answers to the validity assessments, as Cook and Campbell (1979) state "it is unrealistic to expect that a single piece of research will effectively answer all of the validity questions surrounding even simple causal relationship".

Future research should concentrate on rigorous evaluations of effectiveness of interventions in different settings. Such research possibly with non-volunteer participants and where practitioners have lower levels of compliance may produce different conclusions to the study findings. In dentistry, the knowledge of altering professional practice and particularly as it relates to the effective utilisation of clinical guidelines is somewhat limited (McGlone *et al* 2001). There is an urgent need for further research in the area of altering professional practice in dentistry, which will ensure efficient and effective use of limited resources, with the potential to improve the delivery of care by promoting best practice (McGlone *et al* 2001). To achieve this, a research culture in dentistry should be advocated and supported and in future a greater participation by diverse dental practices in such rigorous trials should be encouraged and rewarded.

It is almost inevitable for trials to evaluate the effectiveness of interventions to alter professional practice to undertake a cluster randomisation and this has significant implications for sample size calculations. In order for a study to have adequate power, the sample size should be adjusted for the likely effect of cluster randomisation. A

major problem for calculating sample sizes for cluster randomisation is estimating the likely intraclass correlation. The present study provides useful information on the intraclass correlation that should facilitate future sample size calculations of such trials.

In conclusion there was no evidence that the implementations of CAL-DS and A&F, independently, were successful in changing the general dental practitioners' behaviour to increase their compliance with the SIGN guideline for the management of impacted and unerupted third molar teeth as compared with mailing and the opportunity to attend a postgraduate continuing education course. Related research suggests that these interventions may act as reinforcement of the guideline messages in a complex way. They managed to influence dentists' cognitions even if they did not alter post-guideline behaviour (Bonetti *et al* 2003).

References

1. Office of Technology Assessment of the Congress of the United States. Assessing the efficacy and safety of medical technologies. Washington DC: US Government Printing Office 1978.

2. Department of Health, Working for patients, paper No 6, medical audit, London: HMSO 1989.

3. Institute of Medicine, Guidelines for clinical practice: from development to use. Washington DC, National Academic Press 1992.

4. US Agency for Health Care Policy and Research (AHCPR). Department of Health and Human Services. Acute pain management: operative or medical procedures and trauma. Rockville (MD): The Agency; Clinical Practice Guideline No.1. AHCPR Publication 1993, No. 92-0023.

5. The concise Oxford dictionary. Seventh ed. 1997, Oxford, Oxford University Press.

6. The Cochrane Collaboration Oral Health Group. Newsletter 1999; (4).

7. Agency for Health Care Research and Quality (AHCRQ). Available: http://www.ahcpr.gov. 2000.

8. NHS Centre For Reviews and Dissemination. Available: http://www.york.ac.uk/inst/crd/centre.htm. 2000.

9. General Dental Council (GDC). Available: http://www.gdc-uk.org. 2000.

10. The organizing Medical Network Information (OMNI). Available: http://www.omni, ac.uk. 2000.

11. The Cochrane Collaboration, Cochrane Brochure. Available: hiru.macmaster.ca/cochrane. 2002.

12. The National Institute for Clinical Excellence (NICE). Available: http://www.nice.org.uk. 2002.

13. Absi E, Shepherd J. A comparison of morbidity following the removal of lower third molars by the lingual split and surgical bur methods. International Journal of Oral and Maxillofacial Surgery 1993; 22: 149-153.

14. Ahlqwist M, Grondahl H. Prevalence of impacted teeth and associated pathology in middle-aged and older Swedish women. Community Dental Oral Epidemiology 1991; 19:116-119.

15. Alderson P. The Cochrane Collaboration: An introduction. Evidence Based Dentistry 1998; 1(1): 25-26.

16. Alling C. Dysesthesia of the lingual and inferior alveolar nerves following third molar surgery. Journal of Oral and Maxillofacial Surgery 1986; 44: 454-457.

17. Alling C, Catone G. Management of impacted teeth. Journal of Oral and Maxillofacial Surgery 1993; 51(Supplement 1): 3-6.

18. Antezak-Bouckoms A, Shaw W. The Cochrane Collaboration: Oral Health Group. Journal of Dental Research 1994; 73[11]: 1674-1676.

19. Bader J, Shugars D. Variations, treatment outcomes, and practice guidelines in dental practice. Journal of Dental Education 1995; 59: 61-95.

20. Bader J, Ismail A, Clarkson J. Evidence-based dentistry and the dental research community. Journal of Dental Research 1999; 78(9): 1480-1483.

21. Balas E, Austin S, Mitchell J. The clinical value of computerised information services: a review of 98 randomised clinical trials. Archives of Family Medicine 1996; 5:271-278.

22. Balas E, Weingarten S, Garb C, Blumenthal D, Boren S, Brown G. Improving preventive care by prompting physicians. Archives of Internal Medicine 2000; 160, 301-308.

23. Batchelor P. The legal and ethical implication of evidence based clinical guidelines for clinicians. Evidence Based Dentistry 2000; 2: 5-6.

24. Bero L, Grilli R, Grimshaw J, Harvey E, Oxman A, Thomson M. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. British Medical Journal 1998; 317:465-468.

25. Birch S, Leake J, Lewis D. Economic issues in the development and use of practice guidelines: an application to resource allocation in dentistry. Community Dental Health 1995; 13:70-75.

26. Black N. The relationship between evaluative research and audit. Journal of Public Health Medicine 1992; 14:361-366.

27. Bland J, Kerry S. Statistics notes: weighted comparison of means. British Medical Journal 1998; 316: 129.

28. Bonetti D, Johnston M, Pitts N, Deery C, Ricketts I, Bahrami M *et al.* Can psychological models bridge the gap between clinical guidelines and clinicians' behaviour? A randomised controlled trial of an intervention to influence dentists' intention to implement evidence-based practice. British Dental Journal 2003; 195: 403-407.

29. Bouchier I. Getting Research into Practice. Health Bulletin 1997; 55(1): 62-67.

30. Brann C, Brickley M, Shepherd J. Factors influencing nerve damage during lower third molar surgery. British Dental Journal 1999; 186(10): 514-516.

31. Brewin C, Bradley C. Patient preferences and randomised clinical trials. British Medical Journal 1989; 299: 684-685.

32. Brickley M, Shepherd J, Mancini G. Comparison of clinical treatment decisions with US National Institutes of Health consensus indications for lower third molar removal. British Dental Journal 1993; 175:102-105.

33. Brickley M, Armstrong R, Shepherd J, Kay E. (I) The relevance of health state utilities to lower third molar surgery. International Dental Journal 1995; 45:124-128.

34. Brickley M, Kay E, Shepherd J, Armstrong R. (II) Decision analysis for lower-third-molar surgery. Medical Decision Making 1995; 15:143-151.

35. Brickley M, Tanner M, Evans D, Edwards M, Armstrong R, Shepherd J. (I) Prevalence of third molars in dental practice attenders aged over 35 years. Community Dental Health 1996; 13:223-227.

36. Brickley M, Shepherd J. An investigation of the rationality of lower third molar removal, based on USA National Institutes of Health criteria. British Dental Journal 1996;180 [7]: 249-254.

37. Brickley M, Shepherd J, Mancini G. (II) Comparison of clinical treatment decisions with US National Institutes of Health consensus indications for lower third molar removal. British Dental Journal 175, 102-105. 1996.

38. British Dental Association. BDA Dental business trends survey. British Dental Association, London 2000;1-12.

39. Brokaw W. The third molar question: when and why should we recommend removal? Virginia Dental Journal 1991; 68, 18-21.

40. Bruce R, Frederickson G, Small G. Age of patients and morbidity associated with mandibular third molar surgery. Journal of the American Dental Association 1980; 101: 240-245.

41. Cabana M, Rand C, Powe N, Wu A, Wilson M, Abboud P *et al.* Why don't physicians follow clinical practice guidelines? A framework for improvement. Journal of the American Medical Association 1999; 282(15): 1458-1465.

42. Campbell M, Grimshaw J. Cluster randomised trials: time for improvement. British Medical Journal 1998; 317, 1171-1172.

43. Campbell M, Mollison J, Steen N, Grimshaw J, Eccles M. Analysis of cluster randomised trials in primary care: a practical approach. Family practice 2000; 17 [2], 192-196.

44. Campbell M, Mollison J, Grimshaw J. Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. Statistics In Medicine 2001; 20: 391-399.

45. Cargill V, Cohen D, Kroenke K, Neuhauser D. Ongoing patient randomisation: an innovation in medical care research. Health Service Research 1986; 21: 663.

46. Carmichael F, McGowan D. Incidence of nerve damage following third molar removal: a West of Scotland Oral Surgery Research Group Study. British Journal of Oral and Maxillofacial Surgery 1992; 30:78-82.

47. Casagrande J, Pike M, Smith P. The power function of the test for comparing two binomial distributions. Applied Statistics 1978; 27: 176-180.

48. Chalmers I, Dickersin K, Chalmers T. Getting to grips with Archie Cochrane's agenda. British Medical Journal 1992; 30: 786-787.

49. Chalmers T, Celano P, Sacks H, Smith H. Bias in treatment assignment in controlled clinical trials. New England Journal Of Medicine 1983; 309:1358-1361.

50. Chestnutt I, Kinane D. Factors influencing the diagnosis and management of periodontal disease by general dental practitioners. British Dental Journal 1997; 183: 319-323.

51. Chiapasco M, De Cicco L, Marrone G. Side effects and complications associated with third molar surgery. Oral Surgery, Oral Medicine and Oral Pathology 1993; 76:412-420.

52. Clarkson J, Worthington H, Chalmers I. Reducing harm and maximising benefit. Evidence Based Dentistry 1999; 1[2]: 4-5.

53. Cochrane A. Effectiveness and efficiency. Random reflections on health services. London: Nuffield Provincial Hospital Trust, 1972.

54. Cook T, Campbell D. Quasi-experimentation: Design and analysis issues for field setting. Boston: Houghton Mifflin, 1979.

55. Crombie I, Davis H, Abraham S. The audit handbook: improving health care through clinical audit. Chichester: John Wiley and Sons, 1993.

56. Curren A, Damm D, Drummond J. Pathologically significant pericoronal lesions in adults: histopathologic evaluation. Journal of Oral and Maxillofacial Surgery 2002; 60: 613-617.

57. Daley T. Third molar prophylactic extraction: a review and analysis of the literature. General Dentistry 1996;(44): 310-320.

58. Davidoff F, Haynes R, Sackett D, Smith R. Evidence based medicine: a new journal to help doctors identify the information they need. British Medical Journal 1995; 310: 1085-1086.

59. Davis D, Thomson O'Brien M, Freemantle N, Wolf F, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education. Do conferences, workshops, rounds, and other traditional continuing medical activities change physician behaviour or health care outcome? Journal of the American Medical Association 1999; 282: 867-874.

60. Davis D, Thomson M, Oxman A, Haynes R. Evidence for the effectiveness of continuing medical education. Journal of the American Medical Association 1992; 268: 1111-1117.

61. Davis D, Thomson O'Brien M, Oxman A, Haynes R. Changing physician performance: a systematic review of the effect of continuing medical education strategies. Journal of the American Medical Association 1995; 274(9): 700-705.

62. Day J, Altman D. Blinding in clinical trials and other studies. British Medical Journal 2000; 321: 504-506.

63. Deery C, Fyffe H, Nuttall N, Nugent Z, Pitts N. The dental caries status of Scottish adolescents reported to be regular attenders. Initial results from a primary dental care based research network. British Dental Journal 1999; 187: 95-99.

64. Department of health. A first class service-quality in the new NHS,UK White Paper. Department of Health, The Stationery Office: London. 1998.

65. Donner A, Birkett N, Buck C. Randomisation by cluster: Sample size requirements and analysis. American Journal of Epidemiology 1981; 114: 906-934.

66. Eccles M, Mason J, Freemantle N. Developing valid cost effectiveness guidelines: a methodological report from the North of England evidence based guideline development project. Quality in Health Care 2000; 9: 127-132.

67. Eccles M, Steen N, Grimshaw J, Thomas L, McNamee P, Soutter J *et al.* Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. Lancet 2001; 357; 1406-1409.

68. Edwards S, Lilford R, Braunholtz D, Jackson J, Hewison J, Thornton J. "Ethics of randomised trials" in Health Services Research methods. A guide to best practice. London: BMJ, 1998.

69. Effective Health Care. Implementing clinical practice guidelines. Can guidelines be used to improve clinical practice? Bulletin 1994; (8): 1-12.

70. Effective Health Care. Getting evidence into practice. Bulletin 1999; 5(1): 1-16.

71. Effectiveness Matter. Prophylactic removal of impacted third molars: is it justified? The University of York 1998; 3(2).

72. Eklund S, Pittman J. Variations in patterns of third molar removal. Journal of Dental Research 2000; 79 (Abs 475).

73. Eklund S, Pittman J. Third molar removal patterns in an insured population. Journal of the American Dental Association 2001; 469-474.

74. Elderton R, Nuttall N. Variation among dentists in planning treatment. British Dental Journal 1983; 154: 201-206.

75. Ellis J, Mulligan I, Rowe J, Sackett D. Inpatient general medicine is evidence based. Lancet 1995; 346: 407-410.

76. Esposito M, Coulthard P, Worthington H, Jokstad A. Quality assessment of randomized controlled trials of oral implants. International Journal of Oral and Maxillofacial Implants 2001;16[6], 783-792.

77. Faculty of General Dental Practitioners (UK). FGDP Guideline: Selection criteria for dental radiography1998.

78. Faculty of General Dental Practitioners (UK). FGDP Guideline: Clinical examination and record keeping good practice 2001.

79. Feder G. Management of mild hypertension: which guidelines to follow? British Medical Journal 1994; 308: 470-471.

80. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J. Using clinical guidelines. British Medical Journal 1999; 318: 728-730.

81. Fletcher R, Fletcher S. Clinical epidemiology: the essentials. 3rd ed. Baltimore, MD: Williams and Wilkins 1996.

82. Freemantle N, Grilli R, Grimshaw J, Oxman A. Implementing findings of medical research: the Cochrane Collaboration on effective professional practice. Quality in Health Care 1995; (4): 45-47.

83. Freemantle N, Harvey E, Wolf F, Grimshaw J, Grilli R, Bero L. Collaboration on effective professional practice module. The Cochrane Database of Systematic Reviews. (Cochrane Review) The Cochrane Library 1999; (2).

84. Freemantle N, Harvey E, Wolf F, Grimshaw J, Grilli R, Bero L. Printed educational materials: effects on professional practice and health care outcomes. (Cochrane Review) The Cochrane Library 2002; (1).

85. Fyffe H. DHSRU at Dentistry 2000 survey. Tuith online 2000; 1-3.

86. Gill P, Dowell A, Neal R, Smith N, Heywood P, Wilson A. Evidence based general practice: a retrospective study of interventions in one training practice. British Medical Journal 1996; 312: 819-821.

87. Gilthrope M, Bedi R. An exploratory study combining hospital episode statistics with socio-demographic variables to examine the access and utilisation of hospital oral surgery services. Community Dental Health 1997; 14(4):209-213.

88. Godlee F. Getting evidence into practice: needs for the right resources and the right organisation. British Medical Journal 1998; 317:6-8.

89. Goldberg M, Memerich A, Marco W. Complications after mandibular third molar surgery: a statistical analysis of 500 consecutive procedures in private practice. Journal of the American Dental Association 1985; 111: 277-279.

90. Goodey R, Brickley M, Hill C, Shepherd J. A controlled trial of three referral methods for patients with third molars. British Dental Journal 2000; 189[10]: 556-560.

91. Grilli R, Lomas J. Evaluating the message: the relationship between compliance rate and the subject of practice guideline. Medical Care 1994; 32: 202-213.

92. Grilli R, Ramsay C, Minozzi S. Mass media interventions: effects on health services utilisation. (Cochrane Review) The Cochrane Library 2002; (1).

93. Grimshaw J, Russell I. (II) Achieving health gain through clinical guidelines. I: Developing scientifically valid guidelines. Quality in Health Care 1993; 2:243-248.

94. Grimshaw J, Russell I. (I) Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993; 342: 1317-1322.

95. Grimshaw J, Russell I. Achieving health gain through clinical guidelines. II: Ensuring guidelines change medical practice. Quality in Health Care 1994; 3: 45-52.

96. Grimshaw J, Freemantle N, Wallace S, Russell I, Hurwitz B, Watt I. Developing and implementing clinical practice guidelines. Quality in Health Care 1995; [4]: 55-64.

97. Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. Family Practice 2000; 17[Supplement 1]: 11-16.

98. Grol R. Implementing guidelines in general practice care. Quality in Health Care 1992; 91:184-191.

99. Grol R. Research and development in quality of care: establishing the research agenda. Quality in Health Care 1996; 5:235-242.

100. Grol R. Personal paper: Beliefs and evidence in changing clinical practice. British Dental Journal 1997; 315:418-421.

101. Grol R, Jones R. Twenty years of implementation research. Family Practice 2000; 17[Suppl.1]: 32-35.

102. Grondahl H. Radiographic caries diagnosis. A study of caries progression and observer performance. Swedish Dental Journal Supplement 1979; 3: 1-32.

103. Guven O, Keskin U, Akal U. The incidence of cysts and tumours around impacted third molars. International Journal of Oral and Maxillofacial Surgery 2000; 29[131]: 135.

104. Haines A, Feder G. Guidance on guidelines: Writing them is easier than making them work. British Medical Journal 1992; 305: 785-786.

105. Haines A, Jones R. Implementing findings of research. British Medical Journal 1994; 308: 1488-1492.

106. Haines A, Donald A. Getting research findings into practice, Making better use of research findings. British Medical Journal 1998; 317:72-75.

107. Harbour R, Miller J. A new system for grading recommendations in evidence based guidelines. British Medical Journal 2001; 323: 334-336.

108. Harradine N, Pearson M, Toth B. The effect of extraction of third molars on late lower incisor crowding: A randomised controlled trial. British Journal of Orthodontics 1998; 25: 117-122.

109. Haycox A, Bagust A, Walley T. Clinical guidelines - the hidden costs. British Medical Journal 1999; 318:391-393.

110. Haynes B, Haines A. Barriers and bridges to evidence based clinical practice. British Medical Journal 1998; 317: 273-276.

111. Haynes R, Davis D, McKibbon K. A critical appraisal of the efficacy of continuing medical education. Journal of the American Medical Association 1984; 251: 61-64.

112. Haynes R. Where's the meat in clinical journals? The American College of Physicians (ACP) Journal Club 1993; 119:23-24.

113. Haynes R, Sackett D, Grey M, Cook D, Guyatt G. Transferring evidence from research into practice: 2. Getting the evidence straight. Evidence Based Medicine 1997; 2(1):4-6.

114. Hill A. Suspended judgement. Memories of the British Streptomycin Trial in Tuberculosis. The first randomised clinical trial. Controlled Clinical Trials 1990; 11: 77-79.

115. Hulscher M, Wensing M, Van der Weijden T, Grol R. Interventions to implement prevention in primary care. (Cochrane Review) The Cochrane Library [1]. 2002.

116. Hunt D, Haynes R, Hanna S. Effects of computer-based clinical decision support systems on physician performance and patient outcomes. A systematic review. Journal of the American Medical Association 1998; 280: 1339-1346.

117. Hunt D, Jaeschke R, McKibbon K. Users' guides to the medical literature. XXI. Using electronic health information resources in Evidence-Based Practice. Journal of the American Medical Association 2000; 283 (14).

118. Hunt R, Kohout F, Beck J. Use of pit and fissure sealants in private dental practices. British Dental Journal 1983; 62: 234-237.

119. Hunter D. Rationing and evidence-based medicine. Journal of Evaluation of Clinical Practice 1996; 2: 5-8.

120. Hurwitz B. Legal and political considerations of clinical practice guidelines. British Medical Journal 1999; 318:661-664.

121. Jadad A. Randomised controlled trials; A user's guide. 1st ed. London: British Medical Journal (BMJ) Publishing Group, 1998.

122. Johnston G, Crombie I, Alder E, Davis H, Millard A. Reviewing audit: barriers and facilitating factors for effective clinical audit. Quality in Health Care 2000; 9: 23-36.

123. Johnston M, Langton K, Haynes R, Mathieu A. The effects of computer based clinical decision support systems on clinician performance and patient outcome. Annals of Internal Medicine 1994; 120:135-142.

124. Kay E, Blinkhorn A. A qualitative investigation of factors governing dentists' treatment philosophies. British Dental Journal 1996; 180: 171-176.

125. Kay E, Nuttall N. Clinical decision making: an art or a science? British Dental Journal (BDJ) books, published by the British Dental Association, London 1997; 1-77.

126. Kay E, Silkstone B, Worthington H. Evaluation of computer aided learning in developing clinical decision- making skills. British Dental Journal 2001; 190[10]: 554-557.

127. Kelly M, Steel J, Nuttall N, Brandneck G, Morris J, Nunn J *et al.* Adult Dental Health survey, oral health in the United Kingdom. Office for National statistics, 1998.

128. Kipp D, Goldstein B, Weiss W. Dysesthesia after mandibular third molar surgery: a retrospective study and analysis of 1377 surgical procedures. Journal of the American Medical Association 1980; 100: 185-192.

129. Knutsson K, Brehmer B, Lysell L, Rohlin M. I .General dental practitioners' evaluation of the need for extraction of asymptomatic mandibular third molars. Community Dental Oral Epidemiology 1992; 20: 347-350.

130. Knutsson K, Brehmer B, Lysell L, Rohlin M. II. Asymptomatic mandibular third molars: oral surgeon's judgement of the need for extraction. Journal of Oral and Maxillofacial Surgery 1992; 50: 329-333.

131. Knutsson K, Lysell L, Rohlin M, Brickley M, Shepherd J. II. Comparison of decisions regarding prophylactic removal of mandibular third molars in Sweden and Wales. British Medical Journal 2001; 190[4]: 198-202.

132. Knutsson K, Lysell L, Rohlin M. I. Dentists' decisions on prophylactic removal of mandibular third molars: a 10-year follow-up study. Community Dental Oral Epidemiology 2001; 29: 308-314.

133. Kostopoulou O, Brickley M, Shepherd J, Knutsson K, Rohlin M. Agreement between practitioners concerning removal of asymptomatic third molars. Community Dental Health 1997; 14: 129-132.

134. Kostopoulou O, Brickley M, Shepherd J, Newcombe R, Knutsson K, Rohlin M. Intra-observer reliability regarding removal of asymptomatic third molars. British Dental Journal 1998; 184: 557-559.

135. Kostopoulou O, Brickley M, Shepherd J, Newcombe R. Perceived risk of future pathology associated with pathology-free third molars: a comparison of oral and maxillofacial surgeons and family dentists. British Dental Journal 2000; 188(1): 28-31.

136. Landes DP. The relationship between dental health and variations in the level of third molar removals experienced by populations. Community Dental Health 1998; 15:67-71.

137. Larsen P. Alveolar osteitis after surgical removal of impacted third molars. Identification of patient at risk. Oral Surgery, Oral Medicine and Oral Pathology 1992; 73: 393-397.

138. Laskin D. Evaluation of the third molar problem. Journal of the American Dental Association 1971; 82: 824-828.

139. Last J. A dictionary of epidemiology. Third ed. Oxford: Oxford University Press, 1995.

140. Lawrence A. Welcome to Evidence Based Dentistry. Evidence Based Dentistry 1998; 1[1]: 2-3.

141. Liddle J, Williamson M, Irwing L. Method for evaluating research and guidelines evidence. Sydney: NSW Health Department 1996.

142. Lomas J. Words without action? The production, dissemination, and impact of consensus recommendations. Annals of Review Public Health 1991; 12: 41-65.

143. Lomas J. II. Teaching old (and not so old) docs new tricks: Effective ways to implement research findings. CHEPA Working Paper Series 1993; (93-4).

144. Lomas J. I. Diffusion, Dissemination, Implementation: Who should do what? Annals of New York Academy of Science 1993; 703: 226-235.

145. Lopes V, Mumenya R, Feinmann C, Harris M. Third molar surgery: an audit of the indications for surgery, post-operative complaints and patient satisfaction. British Journal of Oral and Maxillofacial Surgery 1995; 33: 33-35.

146. Mackay F, Thompson R. First Steps in Audit. British Dental Journal 1991; 292-293.

147. Main P, Lewis D, Hawkins R. A survey of general dentists in Ontario, Part 1: Sealant use and knowledge. Journal of Canadian Dental Association 1997; 63: 543-553.

148. Marriott S, Palmer C. Disseminating healthcare information: getting the message across. Quality in Health Care 2000; 9: 58-62.

149. Mason D. Lingual nerve damage following lower third molar surgery. Journal of Oral and Maxillofacial Surgery 1988; 17: 290-294.

150. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analysis? Lancet 2000; 356: 1228-1231.

151. McColl E, Smith M, Whitworth J, Seccombe G, Steele J. Barriers to improving endodontic care: the views of NHS practitioners. British Dental Journal 1999; 186(11): 564-568.

152. McGlone P, Watt R, Sheiham A. Evidence-based dentistry: an overview of the challenges in changing professional practice. British Dental Journal 2001; 190: 636-639.

153. McKibbon K, Wilczynski N, Walker-Dilks C. How to search for and find evidence about therapy. Evidence Based Medicine 1996; 1: 70-72.

154. Mercier P, Precious D. Risks and benefits of removal of impacted third molars. Journal of Oral and Maxillofacial Surgery 1992; 21: 17-27.

155. Mercier P, Long A, Ralph J, Bailey H. Audit activity and uptake of postgraduate dental education among general dental practitioners in Yorkshire. British Dental Journal 1998; 184(3): 138-142.

156. Mittman B, Tonesk X, Jacobson P. Implementing clinical guidelines: Social influence strategies and practitioner behaviour change. Quality Review Bulletin 1992; 18(12): 413-422.

157. Moher D, Alejandro R, Jadad M, Nichol G, Penman M, Tugwell P *et al.* Assessing the quality of randomised controlled trials: An annotated bibliography of scales and checklists. Controlled Clinical Trials 1994; 16: 62-73.

159. Moher D, Schulz K, Altman D, for the CONSORT Group. I. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel- group randomised trials. Journal of the American Medical Association 2001; 285[15]: 1987-1991.

159. Moher D, Jones A, Lepage L, for the CONSORT Group. II. Use of CONSORT statement and quality of reports of randomised trials: a comparative before-and-after evaluation I. Journal of the American Medical Association 2001; 285: 1992-1995.

160. Moles D, dos Santos Silva I. Causes, associations and evaluating evidence: Can we trust what we read? Evidence Based Dentistry 2000; 2: 75-78.

161. Montenegro R, Needleman I, Moles D, Tonetti M. Quality of RCTs in periodontology, a systematic review. Journal of Dental Research 2002; 81[12]: 866-870.

162. Morgan R. Quality evaluation of clinical records of a group of general dental practitioners entering a quality assurance programme. British Dental Journal 2001; 191[8]: 436-441.

163. Muir Gray J, Haynes R, Sackett D, Cook D, Guyatt G. Transferring evidence from research into practice: 3. Developing evidence-based clinical policy. Evidence Based Medicine 1997; 2[2]: 36-38.

164. Mulrow C. The medical review article: state of the science. Annals of Internal Medicine 1987; 106[485]: 458.

165. Mulrow C. Rationale for systematic reviews. British Medical Journal 1994; 309: 597-599.

166. Murphy E. The logic of medicine. Baltimore: John Hopkins University Press, 1976.

167. Needleman I. Is this good research? Look for CONSORT and QUORUM. Evidence Based Dentistry 2000; 2: 61-62.

168. Newell D. Intention-to-treat analysis: implications for quantitative and qualitative research. International Journal of Epidemiology 1992; 21: 837-841.

169. Newton-Syms F, Dawson P, Cooke J, Feely M, Booth T, Jerwood D. The influence of an academic representative on prescribing by general practitioners. British Journal of Clinical Pharmacology 1992; 33: 69-73.

170. Nitzan D, Keren T, Marmary Y. Does an impacted tooth cause root resorption of the adjacent one? Journal of Oral Surgery 1981; 51, 221-224.

171. O'Brien K, Wright J, Conboy F, Bagley L, Lewis D, Read M *et al.* The effect of orthodontic referral guidelines: a randomised controlled trial. British Dental Journal 2000; 188[7]: 392-396.

172. Ogden G, Bissias E, Ruta D, Ogston S. Quality of life following third molar removal: a patient versus professional perspective. British Dental Journal 1998; 185(8): 407-410.

173. Osborn T, Federickson G, Small I, Torgerson T. A prospective study of complications related to mandibular third molar surgery. Journal of Oral and Maxillofacial Surgery 1985; 43: 767-769.

174. Oxman A, Sackett D, Guyatt G. How to get started. Journal of the American Medical Association 1993; 270: 2093-2095.

175. Oxman A. No magic bullets: A systematic review of 102 trials of interventions to help health care professionals deliver services more effectively or efficiently. Canadian Medical Association Journal 1995; 153: 1423-1431.

176. Oxman A, Flottrop S. An overview of strategies to promote implementation of evidence based health care. In siagy C and Haines A (eds) Evidence based practice in primary care. London: British Medical Journal (BMJ) Books 1998.

177. Oxman A. Cochrane Effective Practice and Organisation of Care Group (EPOC). Available in : http://www.abdn.ac.uk/public_health/hsru/epocsc.htm. 1999.

178. Petrie J, Grimshaw J, Bryson A. The Scottish Intercollegiate Guidelines Network Initiative: Getting validated guidelines into local practice. Health Bulletin 1995; 53(6): 345-348.

179. Pitts N, Nugent Z, Bedi R. Mirroring the Scottish Dental Practice Board Database to facilitate Health Services Research. Journal of Dental Research 1997; 76: 1019.

180. Pocock S. Clinical trials: a practical approach. New York: Wiley 1983.

181. Pratt C, Hekmat M, Barnard J, Zaki G. Indications for third molar surgery. The Royal College of Surgeons of Edinburgh 1998; 43:105-108.

182. Precious D. Summary : No research evidence to support the prophylactic removal of pathology-free impacted third molars. Evidence Based Dentistry 2000; 2: 42-43.

183. Prescott R, Counsell C, Gillespie W, Grant A, Russell I, Kiauka S *et al.* Factors that limit the quality, number and progress of randomised controlled trials. Health Technology Assessment 1999; 3(20).

184. Randall D, Cebul M. Randomised controlled trials using the Metro Firm System. Medical Care 1991; 29[7, Supplement]: 9-18.

185. Rangachari P. Evidence-based medicine: an old French wine with a new Canadian label? Journal of the Royal Society of Medicine 1997; 90: 280-284.

186. Reit C, Grondahl H. Endodontic decision making under uncertainty: A decision analytic approach to management of periapical lesions in endodontically treated teeth. Endodontics and Dental Traumatology 1987; 3: 15-20.

187. Reit C, Grondahl H. Endodontic retreatment decision making among a group of general practitioners. Scandinavian Journal of Dental Research 1988; 96: 112-117.

188. Richards D, Lawrence A. Evidence based dentistry. British Dental Journal 1995; 179[7]: 270-273.

189. Richards D. Review on "RCTs in periodontology do not meet current recommendations on quality". Evidence Based Dentistry 2003; 4: 3.

190. Robinson P. The impacted wisdom tooth: to remove or to leave alone? Dental Update 1994; 21: 245-248.

191. Robinson P, Smith K. Lingual nerve damage during lower third molar removal: a comparison of two surgical methods. British Dental Journal 1996; 180: 456-461.

192. Roland M, Torgerson D. Understanding controlled trials: What are pragmatic trials? British Medical Journal 1998; 316: 285.

193. Rood J. Degrees of injury to the inferior alveolar nerve sustained during the removal of impacted mandibular third molars by the lingual split technique. British Journal of Oral Surgery 1983; 21: 103-116.

194. Rood J. Permanent damage to inferior alveolar and lingual nerves during the removal of impacted mandibular third molars. Comparison of two methods of bone removal. British Dental Journal 1992; 172: 108-110.

195. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. British Dental Journal 1995; 310: 1122-1126.

196. Rud J. The split bone technique for removal of impacted mandibular third molars. Journal of Oral Surgery 1970; 28:416-421.

197. Rushton V, Horner K, Worthington H. Factors influencing frequency of bitewing radiography in general dental practice. British Dental Journal 1996; 24: 272-276.

198. Russell E, Harrold T, Watson S. A Scottish survey of general and community dental practitioners 2000; The Toothousand project. Department of Public Health, University of Aberdeen Medical School 2000.

199. Russell I, Grimshaw J, Wilson B. Scientific and methodological issues in quality assurance. In Becks JS, Bouchier IAD, Russell IT Eds. Quality Assurance in Medical Care. Proceedings of the Royal Society of Edinburgh 1993; 101B: 77-103.

200. Russell I, Grimshaw J. The effectiveness of referral guidelines: a review of the methods and findings of published evaluations. In Coulter A and Roland M Eds. Referrals from General Practice. Oxford: Oxford University Press 1992.

201. Sackett D, Gent M. Controversy in counting and attributing events in clinical trials. New England Journal Of Medicine 1979; 301: 1410-1412.

202. Sackett D. Bias in analytic research. Journal of Chronic Disease 1979; 32: 51-63.

203. Sackett D, Rosenberg W, Grey M, Haynes R, Richardson W. Evidence based medicine: what it is and what it isn't. British Medical Journal 1996; 312:71-72.

204. Sackett D, Richardson W, Rosenberg W, Haynes R. Evidence based medicine: how to practice and teach EBM. London: Churchill Livingstone 1997.

205. Scally G, Donaldson L. Clinical governance and the drive for quality improvement in the new NHS in England. British Medical Journal 1998; 317: 61-65.

206. Schersten E, Lysell L, Rohlin M. Prevelance of impacted third molars in dental students. Swedish Dental Journal 1989; 13: 7-13.

207. Schulz K, Chalmers I, Hayes R, Altman D. Empirical evidence of bias. Journal of the American Medical Association 1995; 273: 408-412.

208. Scottish Dental Practice Board. Dental Practice Board Annual report, Trinity Park House: Edinburgh 1997.

209. Scottish Intercollegiate Guideline Network (SIGN). Preventing dental caries in children at high caries risk. Targeted caries prevention in 6-16 year olds attending for dental care. A national clinical guideline 2000.

210. Scottish Intercollegiate Guidelines Network (SIGN). Management of unerupted and impacted third molar teeth. A national clinical guideline, SIGN 43, 2000.

211. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook, SIGN: Edinburgh. Available: http://www.sign.co.uk. 2002.

212. Secretary of State for Health. The new NHS: modern and dependable. London: Department of Health 1997.

213. Shaneyfelt T, Mayo-Smith M, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer- reviewed medical literature. Journal of the American Medical Association 1999; 281(20): 1900-1905.

214. Shekelle P, Woolf S, Eccles M, Grimshaw J. Developing guidelines. British Medical Journal 1999; 318: 593-596.

215. Shelden T, Chalmers I. The UK Cochrane Centre and The NHS Centre For Reviews and Dissemination: Respective roles within the information systems strategy of the NHS R&D programme, coordination and principles underlying collaboration. Health Economics 1994; 3: 201-203.

216. Shelden T, Guyatt G, Haines A. Getting research findings into practice: When to act on the evidence. British Dental Journal 1998; 317: 139-142.

217. Shepherd J. The third molar epidemic. British Dental Journal 1993; 174: 85.

218. Shepherd J, Brickley M. Surgical removal of third molars. British Medical Journal 1994; 309: 620-621.

219. Sisk A, Hammer W, Shelton D, Joy E. Complications following removal of impacted third molars: the role of the experience of the surgeon. Journal of Oral and Maxillofacial Surgery 1986; 44:855-858.

220. Smart R. The importance of negative results in psychological research. Canadian Journal of Psychology 1964; 5: 225-232.

221. Smith R. Information in practice: What clinical information do doctors need? British Medical Journal 1996; 313: 1062-1068.

222. Smith R. The scientific basis of health services. British Medical Journal 1995; 311:961-962.

223. Song F, Landes D, Glenny A, Sheldon T. Prophylactic removal of impacted third molars: an assessment of published reviews. British Dental Journal 1997; 182:339-346.

224. Soumerai S, McLaughlin T, Avon J. Improving drug prescribing in primary care; a critical analysis of the experimental literature. Milbank Quarterly 1989; 67: 268-317.

225. Stanley H, Alatter M, Collett W, Stringfellow H, Spiegel E. Pathological sequelae of neglected impacted third molar. Journal of Oral Pathology 1988; 17: 113-117.

226. Stephens R, Kogon S, Reid J. The unerupted or impacted third molar- a critical appraisal of its pathologic potential. Journal of Canadian Dental Association 1989; 55: 201-207.

227. Sterne J, Egger M, Smith G. Investigating and dealing with publication and other biases in meta-analysis. British Medical Journal 2001; 323: 101-105.

228. Straus S, Sackett D. Getting research findings into practice. British Dental Journal 1998; 317: 339-342.

229. Sutherland S. The building blocks of evidence-based dentistry. Journal of Canadian Dental Association 2000; 66(5): 241-244.

230. Tate T. Impactions: observe or treat. Journal of California Dental Association 1994; 22: 59-64.

231. The National Institute for Clinical Excellence (NICE). Guidance on the removal of wisdom teeth. A national clinical guideline. available: http://www.nice.org.uk/appraisals/wis_guide.htm. 1999.

232. The Royal College of Surgeons of England Faculty of Dental Surgery. Current clinical practice and parameters of care. The management of patients with third molar teeth 1997.

233. Thomson O'Brien M, Oxman A, Davis D, Haynes R, Freemantle N, Harvey E. (III) Educational outreach visits: effects on professional practice and health care outcomes. (Cochrane Review) In: The Cochrane Library 2002; (1).

234. Thomson O'Brien M, Oxman A, Haynes R. (IV) Local opinion leaders: effects on professional practice and health care outcomes. (Cochrane Review) In: The Cochrane Library 2002; (2).

235. Thomson O'Brien M, Oxman A, Davis D, Haynes R, Freemantle N, Harvey E. (II) Audit and feedback versus alternative strategies: effects on professional practice and health care outcome. (Cochrane Review) In: The Cochrane Library 2002; (1).

236. Thomson O'Brien M, Oxman A, Davis D, Haynes R, Freemantle N, Harvey E. (I) Audit and feedback: effects on professional practice and health care outcomes. (Cochrane Review) In: The Cochrane Library 2002; (1).

237. Torgerson D, Sibbald B. What is a patient preference trial? British Medical Journal 1998; 316: 360-365.

238. Torgerson D. Contamination in trials: is cluster randomisation the answer? British Medical Journal 2001; 322: 355-357.

239. Valmaseda-Castellon E, Berini-Aytes L, Gay-Escoda C. Two percent of lower third molar surgical extractions are associated with only temporary nerve damage. Oral Surgery, Oral Medicine and Oral Pathology 2000; 90: 567-573.

240. Van der Linden W, Cleaton-Jones P, Lownie M. Diseases and lesions associated with third molars. Oral Surgery, Oral Medicine and Oral Pathology 1995; 79:142-145.

241. Venta I, Ylipaavalniemi P, Turtola L. Long-term evaluation of estimates of need for third molar removal. Journal of Oral and Maxillofacial Surgery 2000; 58: 288-291.

242. Von-Wowern N, Nielsen H. The fate of impacted lower third molars after age of 20. International Journal of Oral and Maxillofacial Surgery 1989; 18:277-280.

243. Von Arx D, Simpson M. The effect of dexamethasone on neuropraxia following third molar surgery. British Journal of Oral and Maxillofacial Surgery 1989; 27: 477-480.

244. Wensing M, Van der Weijden T, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective. British Journal of General Practice 1998; 48: 991-997.

245. Wilson S, Delaney B, Roalfe A, Roberts L, Redman V, Wearn M *et al.* Randomised controlled trials in primary care: case study. British Medical Journal 2000; 321:24-27.

246. Woods K. Mega-trials and management of acute mycardial infarction. Lancet 1995; 346: 611-614.

247. Woolf S. Manual for conducting systematic reviews. Agency for Health Care Policy and Research 1996.

249. Woolf S, Grol R, Hutchinson A, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. British Medical Journal 1999; 318:527-530.

250. Woolf S, Rothemich S, Johnson R, Marsland M. Selection bias from requiring patients to give consent to examine data for health services research. Arch Family Medicine 2000; 9: 1111-1118.

251. Woolf S. Shared decision-making: the case for letting patients decide which choice is best. Family Practice 1997; 45: 205-208.

252. Worrall S, Riden K, Haskell R, Corrigan A. UK National Third Molar project: the initial report. Journal of Oral and Maxillofacial Surgery 1998; 36:14-18.

253. Zelen M. A new design for randomised clinical trials. New England Journal Of Medicine 1979; 300: 1242-1245.

Appendix I

A: Project Report

Reference

R2-64

NHS R&D FUNDED PROJECT

EFFECTIVE PRACTICE?: A RANDOMISED TRIAL OF DISSEMINATION AND IMPLEMENTATION STRATEGIES FOR GUIDELINES FOR THE APPROPRIATE EXTRACTION OF THIRD MOLAR TEETH

Final Report

Table of Contents

Executive Summary

Aim of the research

The primary aims of the project were

- to evaluate effectiveness of different dissemination and implementation strategies for evidence based guidelines for the selection of cases for the appropriate management of third molars.
- to undertake the economic evaluation of the development, dissemination and implementation of the guidelines.
- to investigate factors mediating the effects on evidence based practice of two methods of implementing third molar management.
- to obtain values or utilities for individual attributes of third molar care.

Achievements of the research

This randomised trial evaluated two implementation interventions shown from research in primary medical care settings to have greatest potential which were designed to increase the adherence to a SIGN guideline for the appropriate management of third molars by Scottish dentists. The study produced evidence of good adherence with the guideline recommendations at baseline. The self-selecting sample of 51 dentists was randomly allocated to one of four groups according to the protocol. All groups received a copy of the guideline and had the opportunity to attend postgraduate continuing education sessions. The dental practices randomly allocated to an active intervention received either audit and feedback (A&F) or computer assisted learning (CAL) or both. The evaluation did not detect a statistically significant difference between the effectiveness of the implementation interventions and the guideline adherence of the non-intervention group and there was no statistically significant effect of either CAL or A&F on the implementation of these guidelines. Although this study was unable to

detect if CAL or A&F has had any effect in increasing the general dental practitioners compliance with the guidelines , it showed that they may act in a complex way as reinforcement of the guideline messages. The study does provide evidence to support the incorporation of psychological models in the development of trial methodology as well as individual interventions relating to the implementation of dental guideline recommendations.

A statistically significant reduction was detected between the proportion of wisdom tooth extractions at the pre-intervention (37%) and post-intervention period (27%) over all groups. Data from NHS Scotland's Management Information and Dental Accounting System (MIDAS) shows a concomitant reduction in both surgical and non-surgical wisdom tooth extractions took place in Scottish General Dental Practices during the experimental period.

Whilst a reduction in the proportion of third molar extractions by the sample population of dentists was observed, simultaneously, a marked increase in the number of people treated for third molar pathology was measured simultaneously. There are two possible explanations for this phenomenon: that dentists were more confident in their management of the condition as a result of the guideline recommendations or that dentists participating in the trial exhibited improved note-keeping as a consequence of their participation in the trial.

The cost effectiveness analysis detected differences in the costs of the interventions. But they were only statistically significantly different between A&F and the non-intervention groups when compared to the CAL group. This is largely due to the high cost of providing twenty-four dentists randomised to CAL with laptop PCs. A sensitivity analysis of CAL costs where the PC costs are equal to zero failed to demonstrate a statistical difference in costs between the four groups. As the clinical efficacy data did not demonstrate a difference between the intervention groups a cost-minimization strategy is appropriate. The lowest costs were associated with the postgraduate continuing education (PGCE) group.

The trial used psychological models to derive possible mediating factors likely to increase the performing of a behaviour and factors associated with the practice of EBM per se. Dentists' cognitions were demonstrably affected by the interventions with those receiving the A&F intervention increasing their third molar-related knowledge significantly more than dentists who did not receive A&F. Dentists in both A&F and CAL demonstrated a significantly higher intention to extract third molars than their non- intervention group counter-parts but neither of these cognitions were related to evidence based practice. Evidence about the role of psychological models in implementation research is scarce. The identification of beliefs about behaviour as a proxy outcome measure for dental EBM represents an important development in dental implementation research. The findings from the psychological research are consistent with the main trial findings.

Patient preferences for various attributes of third molar treatment were estimated using an independent discrete choice experiment. The relatively small sample size yielded estimates of patients' willingness to pay for appropriate treatment.

Progress of the research

The progress of the research project was as anticipated and outlined in the original proposal with the exception of the withdrawal of five dentists from the study after randomisation. The reasons given for withdrawal were too busy, not interested, moving practice, refurbishment of practice and a change of mind. There were no protocol deviations. The delay in the production of the SIGN guideline resulted in a delay of four to six months in the completion of the project.

The estimate of 40% adherence to the guideline at baseline was based on an estimate of the numbers of extracted wisdom teeth as opposed to the over all management of third molar pathology. The projected estimate of inappropriate third molar extraction was originally set at 33% – 66%. Conversely, the primary outcome measure of interest in the current trial was dentists' adherence to guideline recommendations (appropriate

management of third molars) and the sample population's guideline adherence proved to be substantially higher than expected.

Further research

There is a real need to assess the behavioural, psychological and environmental factors which influence general dental practitioner's participation in research since demand for high quality research in primary dental care is growing, especially for evidence from multi-centre randomised controlled trials. These studies are complex and expensive to conduct. Failure to recruit enough participants from the full range of the general dental practice population may have reduced the generalisability of the trial conclusions.

Our study was unable to detect differences in guideline adherence between the dentists randomised to A&F and CAL&DS. Use of more intensive methods (e.g. educational detailing) to encourage adherence to the guideline might have increased guideline adherence by general dental practitioners but this would have required additional resources and was therefore advised against based on results available from medicine when the grant application was submitted.

There is a need to replicate implementation studies to test for an interaction between the interventions and the generalisability of the interventions into other settings. The high adherence to guidelines at baseline and the lack of variation in the volunteer dentists adherence with the guideline recommendations indicates the need for a larger sample of participants than was initially estimated.

The role of modelling relevant psychological proxy measures to predict behaviour change needs to be investigated in future implementation research.

Importance to NHS and possible implementation

At the pre-intervention stage, the practitioners' overall compliance with the published guidelines was found to be high. It is possible that the use of these interventions in different health technologies, where practitioners have lower levels of compliance might produce different conclusions to these. It appears that the implementation strategies were ineffective in increasing the compliance of the volunteer dental practitioners and this might be a "ceiling" effect as result of this high baseline compliance. Hence greater

participation from diverse dental practices in trials should be encouraged and rewarded in future.

Publication and dissemination

The DHSRU disseminates research findings using a twin-track approach. We have submitted an abstract to disseminate the results in"7th European forum on Quality improvement in health care which will be held in EICC, Edinburgh in 21- 23 March 2002. In addition we will be submitting manuscripts to peer reviewed journals outlining the various multidisciplinary studies embodied in this work .

Publications to date

McIntosh E & Ryan M (1998) The wisdom of using conjoint analysis to estimate utility scores in dental health: the case of third molars. Paper presented at the Health Economists Study Group, The National University of Ireland, Galway 13-14th July 1998.

Ricketts IW, Rowe GWA & Ramsay CD. Clinical Informatics – Computer Based Decision Support for the Implementation of Clinical Guidelines for the Extraction of Third Molars. Paper presented at HC99 held in Harrogate, UK 22-24th March 1999.

Scientists get to root of a teething problem. Scottish Herald. 11 April 1999.

Words of wisdom for dentists. Dundee Evening Telegraph. 12 April 1999.

Computer wisdom can help dentists. Dundee Courier. 13 April 1999.

Third Molars – when to extract. Scottish Dentist. Issue No 39. July/August 1999.

Ramsay C, Ricketts I, Pitts NB, Deery C, Johnston M, McIntosh E, Rennie J. Use of multimedia to inform patient consent prior to 3rd molar extraction. British Computer Society – Medical Special Interest Group 1999 – Video Conference between Glasgow, Swansea, Ulster and Dundee: 1999.

Implementation of Clinical Guidelines – Third Molar Removal as a model. Public Health Seminar, Department of Epidemiology and Public Health, Ninewells Hospital & Medical School, 29 February 2000.

The interesting, The Information & The Controversial. Research in Primary Dental Care. Stirling, 16 March 2000.

Bahrami M, Clarkson JE, Deery C, Gray D, Nugent ZJ, Pitts NB. The recording of a medical history in general dental practice. Scottish School of Primary Care. Stirling: 2001.

Bahrami M, Clarkson JE, Deery C, Gray D, Nugent ZJ, Pitts NB. The appropriateness of third molar management in primary dental care. BSDR, Dublin: 2001.

M, Johnston & Bonetti, D (University of St Andrews); Pitts, N, Deery, C, Ricketts, I, Bahrami, M, Ramsay, C & Johnston J (University of Dundee). Cognitive-behavioural processed in the dissemination of evidence-based guidelines for dental practice. BABCP Abstract Glasgow 2001.

Bonetti, D & M, Johnston (University of St Andrews); Pitts, N, Deery, C, Ricketts, I, Bahrami, M, Ramsay, C & Johnston J (University of Dundee). Using Health Psychology models in implementation research: Results from dentistry. Health Psychology Conference 2001: 5-8 September 2001, St Andrews

M, Johnston & Bonetti, D (University of St Andrews); Pitts, N, Deery, C, Ricketts, I, Bahrami, M, Ramsay, C & Johnston J (University of Dundee). Can implementation intention interventions be used to eliminate a behaviour? Health Psychology Conference 2001: 5-8 September 2001, St Andrews

Bonetti, D & M, Johnston (University of St Andrews); Pitts, N, Deery, C, Ricketts, I, Bahrami, M, Ramsay, C & Johnston J (University of Dundee). Can Psychological models bridge the gap between clinical guidelines and clinicians' behaviour: A

randomised controlled trial of an intervention to influence the intention to implement evidence-based practice. Submitted to Community Dental Health, October 2001

Bonetti, D[1] & M, Johnston[1], Pitts, N[2], Deery, C[3], Ricketts. I[2], Johnston, J[2], Hart, J[1]. Bahrami, M[2] ([1] University of St Andrews, [2] University of Dundee. [3] Edinburgh Dental Institute). A theoretical analysis of identified barriers and facilitators to third molar extraction: What dentists would, should, could and intend to do. Submitted to Community Dental Health, December 2001

McIntosh, E and Louviere J. An Empirical Comparison of Discrete Choice Experiments and Best Attribute Scaling: An application to third molar. Paper accepted for oral presentation at the University of Southern Denmark, 18-19 April, Expert Workshop on Discrete Choice Experiments

Bahrami M. Implementation of guidelines in primary dental care. How effective is it? The Scottish Dental Practice Based Research Network (SDPBRN), 1st annual national symposium, Nov 2001, Royal College of Edinburgh

Effectiveness of Different Dissemination and Implementation Strategies for Evidence Based Guidelines for Selected Cases in Primary Care

Background

An evidence-based approach to clinical practice is advocated to improve the quality of patient care, however there is often a gap between research findings and clinical practice. To address this deficiency there is a need to assist clinicians in access to and adoption of research findings. One possible method of facilitating change in practice is clinical guidelines. It has been shown in medicine that a change in clinical practice in favour of published guidelines is dependent on an active implementation strategy. Consistently effective implementation strategies have not been identified in either medicine or dentistry.

Aim

To investigate the effectiveness of different implementation strategies for evidence based guideline, using the Scottish Intercollegiate Guidelines Network (SIGN) guidelines for appropriate removal of third molar teeth.

Design

A randomised-controlled trial employing a 2x2 factorial design linked to multidisciplinary evaluation.

Subjects

51 volunteer dental practices across Scotland.

Method

Practices were randomly allocated to one of four groups. Pre-intervention data were collected from 49 dental practices. The clinical records of all 16-24 year old patients who attended the practice over a four-month period (August - December 1999) were searched by a clinical researcher who was blind to randomisation. The data extracted included the reason for their attendance and treatment received. This process was repeated following publication of the SIGN Guidelines in April 2000. The post-intervention phase of the project took place between June and October 2000. And data were collected from 46 practices.

Interventions

Mailing of guideline (M); Audit and feedback (A&F); Computer-aided learning with decision support (CAL), and A&F together with CAL. In addition all practitioners had an opportunity to attend a post-graduate course on the guidelines. Thus the M group mirrored current practice to the in dissemination and implementation of the Sign guidelines in primary dental care.

Outcome Measurement

The principle outcome was adherence to the guideline as assessed independently by two researchers. Any disagreement between these evaluators was discussed and an agreement reached.

Results

The overall recruitment rate of practices was 11% of those invited to take part (63 of 565) and this decreased to 8% following the intervention. Prior to the intervention the percent of patients with a problem with third molar teeth was 7% compared with 22% after intervention. This occurred at the same time as a reduction in the overall number of patients seen by the practices (3342 compared with 1935). A statistically significant reduction in the number of patients treated with extraction was detected between the pre- (37%) and post-intervention (27%) phase of this study, (P=0.02). Compliance with

the guidelines was 74% of patients pre-intervention and this increased to 78% post-intervention. However, this difference was not statistically significant (P=0.25). The weighted t-test for audit versus no audit (P=0.62) and CAL versus no CAL was not significant (P=0.76). From the multilevel analysis the odds ratio of compliance with guidelines for dentists who experienced audit versus those who did not was 1.28 (95% CI 0.62 to 2.63) and this compares with an odds ratio of 0.84 (95% CI 0.88 to 1.74) for the CAL dentists versus no CAL. For neither was the difference statistically significant. The study was not sufficiently powered to detect an interaction effect so analyses of the main effects only were undertaken. As there was a weak correlation between pre and post cluster level compliance rates (Product Moment Correlation = -0.125, t = 0.81, n = 43, P>0.4). Therefore all analysis were performed on post intervention compliance rate. All analysis were carried out on intention to treat basis.

Conclusion

There was no statistically significant effect of either CAL&DC or A&F on implementation of these guidelines. This study was unable to show if the CAL&DC and A&F independently had any effect in increasing the general dental practitioners compliance with the guideline but it may act as reinforcement of the guideline messages.

Behavioural Change Study: Using Health Psychology models in dental implementation research

Background

This study used the Theory of Planned Behaviour and Social Cognitive Theory to evaluate the effectiveness of different methods of implementing third molar (TM) management guidelines.

Method

Design: A randomised controlled trial examined the effects of the two implementation methods (2x2 factorial design) on social cognitive variables.

Participants: 51 dentists (41m, 10f), age=42.33 yrs (sd=7.76 yrs).

Measures: ***Cognitions:*** Attitude; Subjective Norm; Perceived Behavioural Control; Intention; Self-efficacy; and Knowledge. ***Evidence-based Practice (EBP) for third molar management*** *TM-EBP* for each dentist was the total number of third molar-related actions (extracting, referring, monitoring, and restoring) judged to be in line with the guidelines they performed, divided by the number of third molar patients they saw, expressed as a percentage. *Extraction-EBP* was the number of third molar- extractions judged to be in line with the guidelines weighted by the number of TM patients.

Procedure: Dentists were randomly allocated to one of four groups - experiencing either, both or neither of two methods: Audit and feedback (A&F) and computer-aided learning with decision support (CAL). All participants completed questionnaires on cognitions relating third molar management before and after the interventions. Patient records and radiographs were examined independently by 2 clinicians, blind to the group allocation, who judged EBP.

Results

1. **Do the interventions influence EBP for third molar management?**

   Neither TM-EBP or Extraction-EBP were significantly influenced by the interventions.

2. **Do the interventions influence cognitions regarding third molar management?**

   A&F increased third molar-related knowledge (Adj. R^2 =.122; $F(1,45) = 4.47$, $p<.05$). there was also an interaction effect of A&F and CAL on intention to extract third molar teeth (Adj. R^2 = .450; $F(1,46) = 4.97$, $p<.05$). A&F and cal increased intention, but experiencing A&F without cal, or cal without A&F, resulted in significantly higher intention compared to the control group.

3. **Can cognitive variables derived from Health Psychology models predict EBP for third molar management?**

   TM-EBP was predicted by dentists' Attitude and Subjective Norm (perceptions of social pressure). Extraction-EBP was predicted by Attitude, Perceived Behavioural Control, and Self-efficacy (confidence in their ability) to do TM extractions.

Conclusion

The interventions influenced Knowledge and Intention, but neither of these variables predicted EBP. Although there were cognitions derived from Health Psychology models which did predict EBP, these variables were not affected by the interventions. This may account for the failure of the interventions to influence EBP. Implementation interventions need a theoretical basis to target cognitions which predict behaviour. Using theoretical models of behaviour change would also enable the use of cognitions as intermediate endpoints, providing a method of evaluating interventions before full trial.

Economic Evaluation of SIGN 3rd Molar Guidelines

Introduction

What's the best way of implementing guidelines? Economists would argue it's the method that allocates resources to their highest valued use. The method is formalised in 'economic evaluation'. This project conducts an economic evaluation of the recent 3rd molar guidelines published by SIGN.

Design

The present value of the costs of development and dissemination of the guidelines and the costs of treating third molar patients were calculated for each of the four implementation groups. A number of assumptions were made in order that scale effects didn't mask true cost differences. In particular we assumed that the implementation strategies applied to all GDPs in Scotland.

The outcome measure used was adherence to the guideline: evidence-based practice (EBP).

Results

Case-mix had important effects on the expected costs of the implementation groups. Patients who were referred, had caries or pericoronitis cost £296, £46 and £37, respectively, more than otherwise identical patients.

The case-mix-adjusted estimates of the baseline cost of the implementation strategies suggested a dichotomy along CAL-DS lines. Specifically, groups with a CAL-DS component were significantly more costly per patient than those groups without a CAL-DS component. Sensitivity analysis ignoring the purchase and distribution costs of PCs attenuated this dichotomy.

The results from the main analysis of EBP found that there was no difference in EBP between any of the implementation groups post-intervention. In an economic evaluation, this simplifies matters considerably. Given no difference between groups post-intervention, the optimal choice of intervention is the least cost intervention: the

intervention in which GDPs received the guideline and additional postgraduate and continuing education (PGCE) courses.

Conclusion

If the measured effects are true and generalisable effects the policy conclusion should be to choose the least costly (PGCE) option.

There are a number of limitations of this study that suggest the effects are neither true nor generalisable. Consequently, the results may reflect the constraints imposed on the design of the study rather than the true cost-effectiveness of the interventions.

Summary of Discrete Choice Experiment Used to Evaluate the Attributes of Third Molar Care

Introduction

The aim of this section of the project was to obtain values for individual attributes of third molar care using discrete choice experiments (DCEs). Any incremental improvements arising in appropriate referral rates arising from the implementation of the third molar guidelines could then attributed a 'value' in terms of increased utility or welfare. Whilst the study showed that none of the implementation strategies gave rise to improved referral rates this sub component of the report serves as a standalone exploration of patient's preferences for attributes of third molar management.

Methods

There are six stages in carrying out the DCE: establishing the attributes; assigning levels to the attributes; devising a statistically efficient design; presenting scenarios; establishing preferences; and analysing the data. The random effects probit model was used to analyse the data, bootstrapping was used to obtain 95% confidence intervals around the coefficients (and resulting welfare values). The inclusion of cost as an attribute in the utility function, allows for the derivation of marginal values for each of the other attributes. Preferences for the DCE scenarios were obtained by surveying 400 members of the general population. A pilot survey was sent to 100 people to test the face validity of the questionnaire.

Results

The attributes of importance in third molar care arising from the literature were as follows: days experiencing dental pain and swelling; episodes of mild pain; prolonged bleeding; nerve damage; crowding of teeth; episodes of painful inflammation of gums; cost of treatment. The response rate for the main survey was 75/400 (19%). Three people failed the consistency checks, giving rise to a consistency rate of 96%. The results show that all attributes of third molar care, apart from nerve damage, are

significant predictors in choice of treatment for third molars. The signs on each of the coefficients are also valid (except mild pain). People are willing to pay £5.91 for a days reduction in severe dental pain, £3.22 for a units reduction in % chance of bleeding requiring a trip to the dentist, £4.75 for a % reduction in probability of crowding of teeth and £9.16 for a reduction in an episode of painful inflammation of the gums.

Discussion

There were a number of limitations with the data obtained from the discrete choice experiment survey, mainly as a result of the poor response rates obtained from the population survey and lack of analysable pilot data to test the specification of the model. As a consequence of this, the results from this stated choice survey cannot claim to be representative of the population's values. However, the results of this discrete choice experiment show that people who do complete such surveys appear to have consistent and rational preferences for dental care, and more specifically have preferences for attributes of third molar care. The results provide information regarding the 'value' patients place on reductions of symptoms of third molar management and have suggested that discrete choice experiments are a potentially useful approach for eliciting values for attributes of dental care where there are clear trade-offs. Future work however should concentrate on strategies to improve the response rates of such survey-based choice experiments.

Conclusion

The discrete choice experiment reported in this study provides a new approach to the elicitation of dental health preferences using an economic preference approach based in random utility theory. Whilst the data obtained from this dental health survey is not representative of the general population due to poor response rates, the results however suggest that this approach is conducive to eliciting consistent preference data and may be a potentially valid way to obtain values for dental health attributes for use in cost-benefit analyses in the future.

Summary of Development of Computer Aided Learning Package(s)

Original Aims

To develop a Computer Aided Learning (CAL) package to provide a multi-media based delivery of the SIGN guidelines for the removal of wisdom teeth.

Methodology

To design and develop the CAL package in a period of 12 months using a proven approach to user-centred software development (Appendix CAL 1-3) incorporating (i) software requirements gathering (ii) software design and implementation (iii) prototyping and evaluation by members of the user group – dentists and patients (Appendix CAL 4-9). A range of multi-media resources were created including videos recorded at Dundee Dental Hospital and School with clinical staff, surgical staff and related experts; electronic X-ray images; and audio 'voice overs'. A range of software tools were used including the multi-media authoring package Asymetrix ToolBook, and various audio, video and graphical editing software and clip art libraries.

Results/Discussion

The resulting software with an online help system is available on CD-ROM complete with installation instructions (Appendix CAL 10-12). An additional achievement was the development of a Patient Information system to provide a multi-media guide for the patient on the known risks and complications for the surgical removal of wisdom teeth. The rationale for including the Patient Information System was to provide dentists with the means to deliver the information required by patients prior to them giving their informed consent. Use of CAL has the potential to release time for the dentist which they could then apply to further study of the SIGN guideline.

Importance to NHS

Importance to NHS includes the potential benefits of CAL in supporting Continuing Professional Development, in a self-help format, with potential for commercial exploitation (CD-ROM)[1].

Future research

Future research has already been undertaken investigating the provision of a multi-media tool to aid those suffering dental injection phobia (McGoldrick *et al* 1999) and further opportunities exist for the provision of training for dentists treating special groups, e.g. HIV patients, the homeless, etc. The research team is also exploring support for a trial of the Patient Information System to ascertain its effectiveness in conveying the known risks and complications associated with the surgical removal of wisdom teeth.

Dissemination of the work includes presentations at the following conferences and meetings:

Applications of Computing in Dentistry, IW Ricketts, C Ramsay, NB Pitts, P McGoldrick, Conference of the Consumer Health Informatics Network, Glasgow, 25th October 2001.

The Use of Multimedia to Inform Patient Consent Prior to 3rd Molar Extraction, C Ramsay, IW Ricketts, NB Pitts, C Deery, M Johnston, E. McIntosh, J Rennie, delivered at Telemedicine and the Use of Computers by Patients, a Video Conference hosted by the Virtual Institutes for Health Informatics in Belfast, Dundee, Glasgow and Swansea, October 20th 1999.

Applying Computer Based Methods to Reducing Needle Phobia in Dental Patients, P. McGoldrick, J Levitt, A Hunter, C Ramsay, IW Ricketts, delivered at Telemedicine and the Use of Computers by Patients, a Video Conference hosted by the Virtual Institutes for Health Informatics in Belfast, Dundee, Glasgow and Swansea, October 20th 1999.

Clinical Informatics - Computer Based Decision Support for the Implementation of Clinical Guidelines for the Extraction of Third Molars, IW Ricketts, GWA Rowe, CD Ramsay, Proceedings Healthcare Computing '99, Harrogate, 22-34th March 1999.

[1] Copyright material from the American Association for Oral and Maxillofacial Surgery used with permission.

Research Workers

Grant Holders

Professor Nigel Pitts	Dental Health Services Research Unit (DHSRU), University of Dundee
Dr Chris Deery	Edinburgh Dental Institute, Lothian Primary Care Trust University of Edinburgh
Dr James Rennie	Scottish Council for Postgraduate Medical And Dental Education (SCPMDE)
Ms Emma McIntosh	Health Economics Research Centre, University of Oxford
Professor Cam Donaldson	University of Calgary, Canada (Formerly of HERU)
Professor Ian Ricketts	Applied Computing, University of Dundee
Dr Glen Rowe	Applied Computing, University of Dundee
Professor Marie Johnston	Department of Psychology, University of St Andrews
Dr Zoann Nugent	DHSRU, University of Dundee

Project Staff (commencement and finish dates in parentheses)

Miss Jennifer Johnston	Research Secretary, based at DHSRU, University of Dundee (28/9/98 – 31/01/02)
Mr Craig Ramsay	Research Assistant, based at Applied Computing, University of Dundee (1/4/98- 31/3/99)
Dr Jacqueline Deas	Clinical Research Fellow (CRF), based at DHSRU, University of Dundee (25/9/98- 1/6/99)
Dr Darren Gray	Clinical Research Fellow (CRF), based at DHSRU, University of Dundee (21/6/99- 10/9/99)
Dr Maryam Bahrami	Clinical Research Fellow (CRF), based at DHSRU, University of Dundee (3/5/99-31/01/02)
Dr Debbie Bonetti	Research Assistant, Department of Psychology, University of St Andrews (1/10/99-28/09/01)

Dr Colin Tilley	Research Fellow in Health Economics, based at DHSRU, University of Dundee
Dr Jan Clarkson	Senior Lecturer, based at DHSRU, University of Dundee
Mrs Marilyn Laird	Unit Administrator, based at DHSRU, University of Dundee

Acknowledgements

Professor Graham Ogden, Oral Surgery, Dundee Dental Hospital

Craig Ramsay, Statistician, Health Services Research Unit, Aberdeen

Graham McLennan, Statistician, Health Services Research Unit, Aberdeen

Professor Jeremy Grimshaw, Health Services Research Unit, Aberdeen

Dr Marion Campbell, Health Services Research Unit, Aberdeen

Elaine Humphries, Audit Facilitator, Scottish Council for Postgraduate Medical and Dental Education

Dr Iain Hunter, General Dental Practitioner

Dr Andrew Wright, General Dental Practitioner

Dr Jim McCafferty, General Dental Practitioner

Dr Colin Yule, General Dental Practitioner

Professor Trevor Sheldon, University of York

Dental Practitioners in Study

Dr William Andrew, Galashiels
Dr M G Arthur, Wishaw, Lanarkshire
Dr William Bennie, Elgin, Moray
Miss Rose-Marie Blair, Newton Mearns, Glasgow
Dr Chris Bloomer, Newport-on-Tay, Fife
Dr Andrew Boggon, Montrose, Angus
Dr Margaret Brown, Inverurie
Dr D A Chong Kwan, Dunfermline, Fife
Mr James Christie, Glasgow
Mr Annibale Coia, Glasgow
Mr Colin M Crawford, Oban, Argyll
Dr Mary Downie, Callander, Perthshire
Dr Alec Dunlop, Helensburgh
Dr Ron Findlay, Dundee
Dr K G E Forrest, Eyemouth, Berwickshire
Dr Carol Garden, Stonehaven, Aberdeenshire
Dr A Geddes, Kirriemuir, Angus
Dr Crawford G Gordon, Thurso, Caithness
Dr Tim Griffith, Elgin, Moray
Dr John Halliday, Inverness
Ms Aisling M Hanly, Blairgowrie, Perthshire
Dr Mark Harrison, Dumbarton
Ms Fiona Hinshalwood, Glasgow
Dr G Horsfield, Bishopbriggs, Glasgow
Dr I A W Hunter, Dundee
Dr David Irvine, East Kilbride
Dr Malcolm Jones, Nairn, Inverness-shire
Dr G Langley, Station Road, Sanquhar, Dumfries
Dr Alison Macaulay, Whitburn
Dr Charles MacDonald, Alloa
Dr Alasdair MacKenzie, Kirkintilloch
Dr Wendy McCombes, Coupar Angus
Dr Ian McDonald, Glasgow
Dr Craig McLellan, Thurso, Caithness
Mrs Linda McNeill, Airdrie
Mr Donald McNicol, Stenhouse Muir, Falkirk
Dr M A Neilson, Johnstone, Renfrewshire
Dr C A Neville, Hawick, Roxburghshire

Dr John O'Neill, Anstruther
Dr David Ovenstone, Schoolhill, Ellon, Aberdeenshire
Dr I G Richmond, Broughty Ferry, Dundee
Dr David Shaw, Fraserburgh
Dr Terry Simpson, Livingston
Mr Martyn Steed, Aberdeen
Dr David Swinfen, High Blantyre, Glasgow
Dr George Taylor, Glasgow
Dr Gordon Thomson, Arbroath
Dr J Watt, Coatbridge

B: Behaviour Science Paper (BDJ 2003,195,7; 403-407)

CAN PSYCHOLOGICAL MODELS BRIDGE THE GAP BETWEEN CLINICAL GUIDELINES AND CLINICIANS' BEHAVIOUR? A RANDOMISED CONTROLLED TRIAL OF AN INTERVENTION TO INFLUENCE DENTISTS' INTENTION TO IMPLEMENT EVIDENCE-BASED PRACTICE.

Bonetti, D.[1], Johnston, M.[2], Pitts, N.[3], Deery, C.[4], Ricketts, I.[5], Bahrami, M.[6], Ramsay, C.[7], and Johnston, J.[8]

[1] Research Fellow, University of St Andrews; School of Psychology, University of St Andrews, St Andrews, KY16 9JU.

[2] Professor of Psychology, University of St Andrews; School of Psychology, University of St Andrews, St Andrews, KY16 9JU.

[3] Director, Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee , DD1 4HR.

[4] Consultant in Paediatric Dentistry / Honorary Senior Lecturer in Paediatric Dentistry Department of Paediatric Dentistry, Edinburgh Dental Institute, Edinburgh.

[5] Professor of Assistive Systems & Healthcare Computing, Department of Applied Computing, University of Dundee, Perth Road, Dundee.

[6] Clinical Research Fellow, Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee , DD1 4HR.

[7] Research Assistant, Department of Applied Computing, University of Dundee, Perth Road, Dundee.

[8] Researcher, Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee , DD1 4HR.

Corresponding author: Dr D. Bonetti: Phone: 01334 462095; e-mail: dlb2@st-andrews.ac.uk

Abstract

Objective: The lag between publication of evidence for clinical practice and implementation by clinicians may be decades. Research using psychological models demonstrates that changing intention is very important in changing behaviour. This study examined an intervention (rehearsing alternative actions) to change dentists' intention to implement evidence-based practice (EBP) for third molar (TM) management.

Design. Randomised controlled trial / Postal.

Setting: Community

Subjects and Methods: Dentists were randomly selected from the Scottish Dental Practice Board Register, then randomly allocated to Intervention or Control groups, and sent a questionnaire. The Intervention Group listed management alternatives to TM extraction prior to their TM extraction intention, and the Control Group did not. Based on psychological models for reducing a behaviour's frequency (EBP is weighted against TM extraction), prior listing of alternatives should decrease extraction intention.

Main outcome measure: Intention to extract TMs.

Results: 99 dentists - 70 Males, 29 Females; mean age = 41.42 years (SD = 8.62). The intervention significantly influenced intention to extract TMs, as desired. Despite similar background and knowledge of management alternatives, participants in the Intervention group had significantly lower intention to extract: Control Group Mean (SD) = .39 (1.99); Intervention Group Mean (SD) = -0.78 (1.89); Mean Difference (SE) = 1.17 (0.42); 95% Confidence Interval for the difference = 0.34 to 1.99.

Conclusion: Results suggest this intervention, which successfully influenced a proximal predictor of behaviour pertinent to dental EBP, may result in improved EBP in a service-level trial. Basing implementation interventions and trial methodology on psychological models may effectively bridge the gap between clinical guidelines and practice.

Introduction

Variation in clinical practice is an important source of variance in health outcomes[1]. The purpose of clinical guidelines is to improve patient outcomes by limiting inappropriate variation by outlining evidence-based practice (EBP).[2-3] This evidence-based approach to care will have an increasing impact on everyday dental practice as more guidelines are introduced. However, it is well-documented (and lamented) that the publication of evidence relating to clinical practice, either as individual studies or as guidelines, does not automatically result in implementation by clinicians.[4-5] The lag between the provision of evidence and its implementation by clinicians may be decades.

There have been over 240 controlled trials of interventions to influence the behaviour of health professionals. Implementation interventions tend to be aimed at increasing knowledge or skills and include approaches involving the dissemination of guidelines and educational materials, small group education, courses, and audit and feedback. However, systematic reviews of such interventions have shown that increasing knowledge and skills is usually insufficient to achieve changes in clinical behaviour.[6-10] Yet, expensive implementation interventions continue to be developed and trialled using this unsuccessful paradigm. There is a need both for more effective methods of designing implementation interventions and for more efficient trial methods

Although implementing guidelines often require clinicians to change their behaviour, there is little evidence that psychological models of behaviour change have been applied to the design of implementation interventions. Yet, these models have been successfully used to predict variation in many different behaviours in many different populations.[11-12] They provide a framework showing relationships between psychological variables, such as beliefs, attitudes and intentions, and behaviour. These models have also been used to design interventions which have been successful in changing behaviour in many different populations.[13] One aim of this study

was to explore the feasibility of applying psychological models to the design of interventions relating to the implementation of EBP, which has yet to be determined.

In addition to enlightening the design of implementation interventions, psychological models may also inform implementation trial methodology. Currently the main means of testing the success of implementation interventions in different populations is in resource intensive service-level trials. Psychological theories model relationships between cognitive variables and behaviour. They therefore identify variables that are proximal predictors of behaviour. Thus, the likelihood of a successful trial may be considerably increased by first examining the effect of interventions on a proximal predictor of behaviour in a modelling experiment. It is reasonable to expect that an intervention which influences a proxy outcome will be more likely to influence behaviour in a full trial than an intervention which does not.

An example of a proximal predictor of behaviour is Intention. While not everyone who intends to perform a behaviour will do so, research using psychological models (particularly the Theory of Planned Behaviour) provides ample evidence that intention to perform a behaviour is nevertheless one of the best predictors of actually performing it.[14-15] It would be expected that an implementation intervention which successfully influences behavioural intention in a modelling experiment would be more likely to change evidence-based practice in a full trial than one which did not.

The Scottish Intercollegiate Guidelines Network (SIGN) has recently published evidence-based guidelines relating to the care and management of third molars.[16] The guideline evidence supports the overall reduction of third molar extractions. An implementation intervention relating to the management of third molars would therefore be required to reduce this behaviour. Based on the psychology literature, the likelihood of a successful implementation trial would be increased if it employs an intervention that reduces dentists' intention to perform third molar extractions.

Research using psychological models provides guidance on designing an intention-behaviour intervention. Gollwitzer, Orbell, Sheeran and their colleagues have demonstrated that the likelihood of performing a behaviour can be increased by planning when you intend to perform it.[17-19] Behavioural approaches point to the need to develop alternative behaviours as the most effective method of eliminating a behaviour.[20] We therefore sought to reduce dentists' intention to extract third molars by having them plan alternative behaviours to extracting third molars.

Method

This was a randomised controlled trial. A preliminary power analysis suggested that a minimum sample of 102 dentists be recruited (across 3 groups: 1 Intervention, 2 Control subgroups) were required to detect an effect size of 0.40, alpha = .05, power = .95 (Faul & Erdfelder's (1992) Gpower program). Approximately 6 months following the postal distribution of guidelines on the management of third molars to all Scottish Dentists, 205 dentists were randomly selected from the Scottish Dental Practice Board Register and allocated to a control or intervention group using a random number generator from SPSS. Figure 1 illustrates the Trial Profile.

Each group was mailed a questionnaire that asked participants to describe their background (post-graduate qualifications; number of years they have been in clinical practice) and their third molar-related experience (number of third molar patients seen in the previous year and month; number of third molar extractions personally performed in these periods). All questionnaires included a general knowledge quiz, derived from research findings relating to third molars. The quiz items covered a wide range of areas, answered on a 3-point scale (True (score = 1), False (score = 0), Not sure (score = 0). Example items are: An asymptomatic third molar should not be removed when it is buried and in close relationship with the inferior dental nerve; Dentigerous cyst formation is rare in association with third molars. Although background variables (demographic, third molar experience and general knowledge) were not expected to be influenced by the intervention, the information was collected to establish any baseline group

differences, since these variables may possibly influence third molar management or the effectiveness of the intervention.

The main outcome measure was Intention to extract third molars. This was measured with 3 questionnaire items. Two items were concerned with the dentists' intention to personally extract third molars: "Of all the patients you see in the next month who require a third molar extraction, approximately how many do you intend to perform?" answered on a 4-point scale (none/some/most/all); How likely is it that you will extract a third molar within the next month? answered on a 7-point scale (Unlikely/Likely); and one item concerned with following the SIGN guidelines (which support doing less third molar extractions): "Do you intend to follow the third molar guidelines?" answered on a 7-point scale (Do/Do not). Since the items were answered on different scales, answers were converted to z scores to ensure equal weighting, and then summed to create a single intention total with higher scores reflecting greater intention to extract third molars.

Intervention

The intervention involved asking participants to develop an alternative behaviour plan using an open question: "If a patient reports to you with third-molar related pain and swelling, what alternative treatments to extraction would you consider?" Participants allocated to the Intervention Group were sent a questionnaire that asked this item prior to the intention items. According to the psychology models, this would have the effect of bringing to mind possible methods of treating third molar problems other than extraction before the formulation of an intention to extract. Having alternative behaviours in mind should thereby inhibit this formulation.

In order to ascertain that groups were equivalent in their specific knowledge (i.e. management alternatives to third molar extraction), a random sample of participants in the Control Group were sent a questionnaire which put this item after the intention items and the rest of the Control Group were sent a questionnaire which did not have this item at all

(subgroups A and B, respectively). The questionnaires for participants in all groups were identical except for the placement of this single item.

Results

Data were analysed using SPSSPC.[21] Group differences were investigated using Chi-Square, t-tests and ANOVA (GLM). Relationships between variables were examined using regression analyses.

Participants

99 dentists agreed to participate in the study by returning the questionnaires: 70 males and 29 females, with Mean age = 41.42 years (SD = 8.62 years). 19/99 participants had been qualified less than 10 years, 20/99 had been qualified between 11 and 15 years, 36/99 had been qualified between 16 and 24 years, and 23/99 had been qualified over 25 years. 19/99 participants had a post-graduate qualification. Participants saw, on average, 19 third molar patients in the previous year (ranging from 0 to 120) and 2 patients in the previous month (ranging from 0 to 15), and personally performed 12 third molar extractions in the previous year (ranging from 0 to 75) and 1 in the previous month (ranging from 0 to 13). The mean score on the general third molar knowledge quiz was 65% (11 out of 17 items; ranging from 0/17 to 15/17) (Cronbach alpha = .43). For the intervention item, the mean number of treatment alternatives to third molar extraction was 3. The treatment options listed were: antibiotics (51 / 99), mouthwash (30 / 99), oral hygiene instruction (22 / 99), periodontal therapy (21 / 99), operculectomy (17 / 99), monitoring (13 / 99), removal of an opposing third molar (13 / 99), pain relief (11 / 99), grinding (10 / 99), restoration (3 / 99), and removal of a second molar (3 / 99).

In order to ascertain, as far as possible, the comparability of respondents and non-respondents, information about non-respondents was sought from the Scottish Dental Register. There were no significant differences between the Respondents and Non-Respondents in either gender (χ^2 (1, 206) = 0.06, p = .88) or years qualified (t (1, 196) = 1.44, p = .15).

We were also able to compare the background data of our participants to an independent random sample of Scottish dentists who participated in a study investigating the influence of audit and feedback and computer-assisted learning on third molar management.[22] There were no significant differences (at $p<.05$) between participants in this study and participants in the independent study in background variables (age: $t(1, 140) = -0.63$, $p = .53$; gender: $\chi^2(1, 150) = 1.64$, $p = .24$; years qualified: $t(1, 147) = -1.11$, $p = .27$, or third molar-related experience (patients seen in the previous year: $t(1, 127) = -1.29$, $p = .20$; or third molar extractions performed in the previous year: $t(1, 129) = -1.07$, $p = .29$).

Equivalence of Groups

There were no significant differences (at $p<.05$) between the Control Subgroups in any variable (return rate, background, independent or dependent) and so the subgroups were combined into a single Control Group for all reported analyses (Control Group N=66, Intervention Group N=33).

There was no significant difference in number of returned questionnaires by group (proportion returned: Intervention Group = .485, Control Group = .481; $z = .054$ i.e. < 1.96). There were no significant differences between the Intervention and Control Groups in any background variable (age: $t(1, 89) = 1.63$, $p = .11$; gender: $\chi^2(1, 98) = 2.38$, $p = .30$; years qualified: $t(1, 96) = 1.13$, $p = .26$), third molar-related experience (patients seen last year: $t(1, 88) = -0.06$, $p = .95$; extractions performed last year: $t(1, 87) = -0.29$, $p = .77$; patients seen last month: $t(1, 87) = 0.43$, $p = .67$; extractions performed last month: $t(1, 85) = -0.47$, $p = .64$), general third molar-related knowledge ($t(1, 97) = -1.06$, $p = .29$), or in the number of alternative treatment options listed ($t(1, 61) = -1.49$, $p = .15$).

Effect of the Intervention on Intention

The intervention was successful in influencing Intention to extract third molars. Dentists in the Intervention Group had significantly lower intention to extract than dentists in the Control

Group: Control Group Mean = 0.39 (SD = 1.99); Intervention Group Mean = -0.78 (SD =1.89); Mean Difference = 1.17 (SE = 0.42); 95% Confidence Interval for the difference = 0.34 to 1.99.

Post Hoc Analyses

The measure of Intention had quite low internal reliability (Cronbach alpha = .40). Post Hoc analyses were performed, exploring the effect of the intervention on each of the three intention items. The pattern was the same for each item as for the overall measure, in that the intervention group scored lower than the control group on each Intention item, although, the difference between the groups was significant (at $p<.05$) for only one of the three items (see Table 1).

Although there were no significant differences between the control and intervention groups in regard to background factors, these factors may still influence the effects of the intervention on dentists' intention to extract third molars. A multiple linear regression analysis was performed to investigate this possibility. However, only Intervention Group significantly contributed to the regression equation (at $p<.05$ level) predicting Intention to extract (See Table 2).

Discussion

Research using psychological models provides evidence that suggests that intention to perform a behaviour is the most reliable predictor of implementing that behaviour. Guideline implementation interventions that do not influence intention to implement EBP are therefore unlikely to influence clinical practice. In this trial we used psychological models to develop an intervention, which successfully *influenced the* intention of dentists to implement third molar EBP in the desired direction. As predicted by behavioural models, planning for alternative behaviours had the effect of reducing intention to perform a specific behaviour incompatible with EBP.

It should be particularly noted that our intention intervention did not add information, unlike other guideline implementation interventions. All participants had received the guidelines

before taking part in the study. Participants in the Intervention and the Control groups were able to demonstrate an equal amount of general knowledge relating to EBP. They were also equally familiar with knowledge relating to possible clinical alternatives to extraction. It was therefore neither the existence of guidelines nor level of knowledge that influenced clinician's intention to implement EBP. This finding may help to explain the general lag in the implementation of EBP and the lack of success of interventions based on educational approaches. Information and knowledge per se are just not enough to motivate EBP.

A limitation of the present study was the return rate, which was just under 50%. There was no offer of recompense for participating in this study, which suggests that dentists who did respond were quite motivated. It is possible that a degree of motivation may be required for the success of this particular intervention. There also may be some concern over how representative the participants in the study were of dentists in Scotland. Nevertheless, there was no significant difference in the return rate between the Intervention and Control groups. There were also no significant differences in gender or years registered between respondents and non-respondents, or between the background of participants and an independent sample of Scottish dentists. There is therefore no reason to believe that the response rate or the background of our particular sample of dentists biased the results.

Conclusion

This study examined the effects of an implementation intervention in the form of modelling experiment. The results of this study demonstrate the effectiveness of a theoretically based intervention. The study also complements and extends current findings on implementation interventions in dental practice. It suggests interventions be tested on a proxy outcome for behaviour, derived from theoretical models, as a possible means of increasing the likelihood of success of service-level trials. While caution is warranted in making generalizations about the effect on EBP, the evidence suggests this intervention, which

successfully influenced a proximal predictor of behaviour pertinent to dental EBP, would be worth investigating in a service-level trial to increase dental EBP.

Applying psychological models to the implementation of dental EBP does not mean ignoring the necessity of educating clinicians in prerequisite knowledge or skills. However, there is ample evidence showing that information transfer is simply not enough to implement changes in clinicians' behaviour. We therefore need to take advantage of research using psychological models specifically directed at behaviour change. These models offer a means of identifying possible target variables, both dependent and independent, for guideline implementation interventions. Designing interventions based on these models also means that the methodology relating to the intervention design can be replicated.

However, using psychological models requires a paradigm shift in guideline implementation studies. The implementation of EBP needs to be conceptualised as behaviour, rather than as ignorance or negligence. Basing implementation interventions on psychological models may be an effective way to bridge the gap between clinical guidelines and clinicians' behaviour.

Acknowledgements: We would like to thank everyone involved in the study, particularly the participating dentists, Marilyn Laird, and Louise Cardno. This project was financed by the NHS R&D program.

References

1 Marteau T M, Johnston M. Health professionals: A source of variance in health outcomes. Psychology and Health 1990; 5: 47-58.

2 Woolf S H, Grol R, Hutchinson A, Eccles M, Grimshaw J. The potential benefits, limitations, and harms of clinical guidelines. British Medical Journal 1999; 318: 527-530.

3 Eccles M, Grimshaw J. (Eds) Clinical Guidelines: From conception to use. Oxford: Radcliffe Press, 2000.

4 Petrie J C, Grimshaw J M, Bryson A. The Scottish Intercollegiate Guidelines Network Initiative (SIGN): Getting validated guidelines into local practice. Health Bulletin 1995; 53 (6): 345-348.

5 Oxman A D, Thomson M A, Davis D A. et al. No magic bullets: A systematic review of 102 trials of interventions to improve professional practice. Canadian Medical Association Journal 1995; 153: 1423-1431.

6 Cranney M, Warren E, Barton S, Gardner K, Wallsy T. Why do GPs not implement evidence-based guidelines? A descriptive study. Family Practice 2001; 18 (4): 359-363.

7 Grimshaw J M, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L, Grilli R, Harvey E, Oxman A, O'Brien M A. Changing provider behaviour – An overview of systematic reviews of interventions. Medical Care 2001; 39 (8): II2-II45.

8 McGlone P, Watt R, Sheiham A. Evidence-based dentistry: An overview of the challenges in changing clinical practice. British Dental Journal 2001; 190 (12): 636-639.

9 Wensing M, Van der Weijden T, Grol R. Implementing guidelines and innovations in general practice: which interventions are effective? British Journal of General Practice 1998; 48: 991-997.

10 Freemantle N, Harvey E L, Wolf F, Grimshaw J M, Grilli R, Bero L A. Printed educational materials: Effects on professional practice and health care outcomes (Cochrane Review). In: The Cochrane Library, 3. Oxford: Update Software, 2001.

11 Norman P, Conner M T. The role of social cognition models in predicting attendance at health checks. Psychology and Health 1993; 8: 447-462.

12 Hounsa A M, Godin G, Alihonou E, Valois P. An application of Ajzen's theory of planned behaviour to predict mothers' intention to use oral rehydration therapy in a rural area of Benin. Social Science and Medicine 1993; 37: 253-261.

13 Conner M, Norman P. (Eds.) Predicting Health Behaviour, Buckingham: Open University Press, 1996.

14 Ajzen I. The Theory of Planned Behaviour. Organizational Behaviour and Human Decision Processes 1991; 50: 179-211.

15 Randall D M, Wolff J A. The time-interval in the intention-behaviour relationship: Meta-analysis. British Journal of Social Psychology 1994; 33: 405-418.

16 SIGN. Management of unerupted and impacted third molar teeth. Guideline 43, Royal College of Physicians, Edinburgh, 2000.

17 Gollwitzer P M, Brandstatter V. Implementation intentions and effective goal pursuit. Journal of Personality and Social Psychology 1997; 73: 186-199.

18 Orbell S., Hodgkins S, Sheeran P. Implementation intentions and the theory of planned behaviour. Personality and Social Psychology Bulletin 1997; 23: 945-954.

19 Sheeran P, Orbell S. Inclined abstainers: A problem for predicting health-related behaviour. British Journal of Social Psychology, (In press).

20 Gambrill E. Handbook of Assessment, Intervention, and Evaluation, London: Jossey-Bass 1977.

21 Norusis, M.J. SPSS for Windows. Base System User's Guide, Release 5.0 Chicago: SPSS Inc. Marketing Department 1992.

22 Effective Practice? A Randomised trial of dissemination and implementation strategies for the appropriate extraction of wisdom teeth. Report to NHSR&D, January 2001

Figure 1 Trial Profile

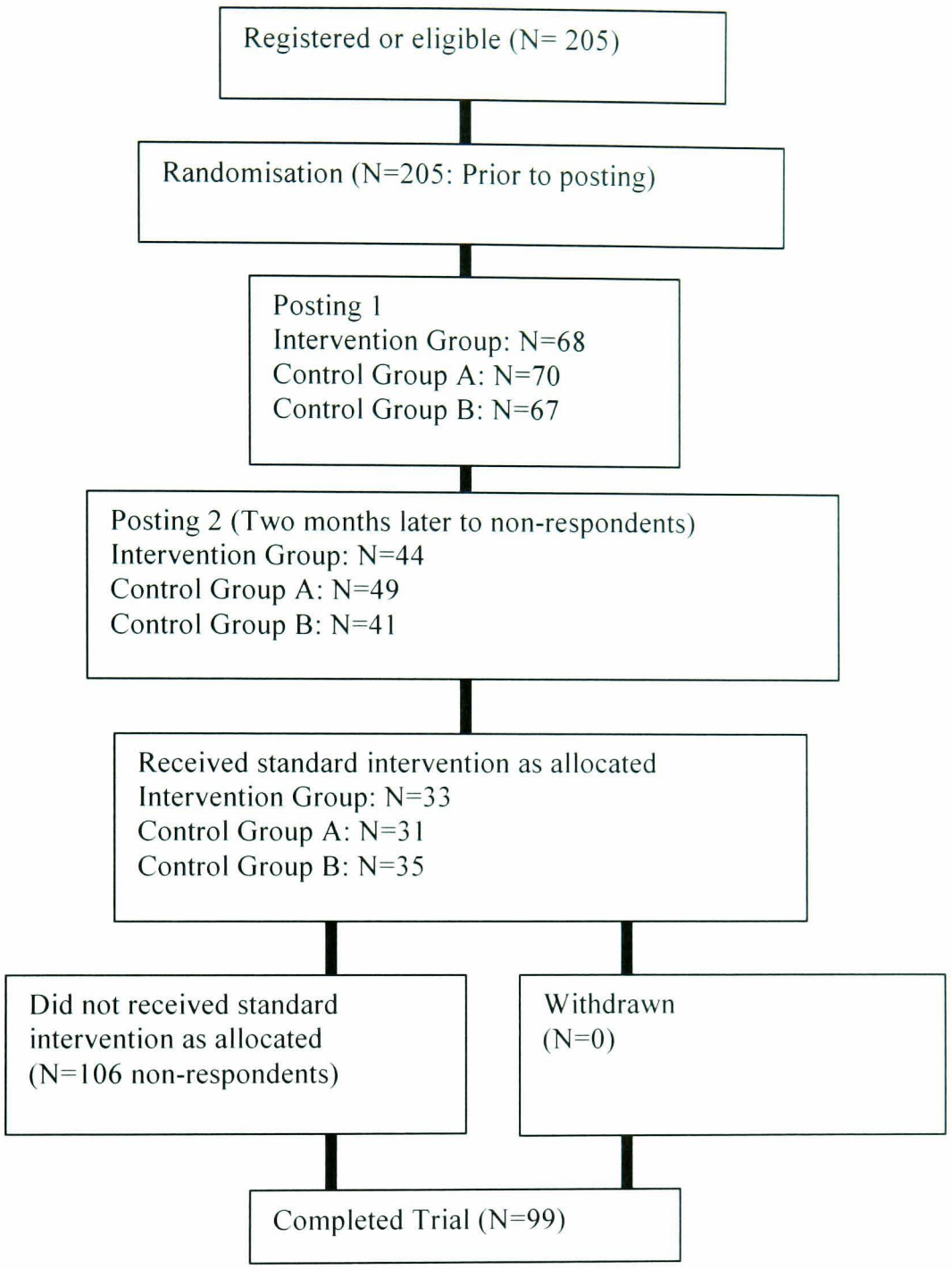


Table 1

Mean Differences between the Control (C) and Intervention (Intv.) Groups for separate Intention items with 95% confidence intervals (CI)

	C Group Mean (SD)	Intv. Group Mean (SD)	Mean Difference	Std. Error	95% CI
Intention item					
1. Of all the patients you see in the next month who require a third molar extraction, approximately how many do you intend to perform?	0.82 (1.02)	-0.16 (0.94)	0.25	0.21	-0.17 to 0.67
2. How likely is it that you will extract a third molar within the next month?	0.10 (0.96)	-0.20 (1.06)	0.32	0.21	-0.10 to 0.74
3. Do you intend to follow the third molar guidelines? [The guidelines support decreasing the number of third molar extractions.]	0.20 (1.07)	-0.40 (0.68)	0.60	0.18	0.25 to 0.95

C = Control; Intv. = Intervention; CI = Confidence Interval

Note: Intention item 3 was reversed scored so that higher scores represent greater intention to extract third molars, as with the other Intention items.

Table 2

Results of the Multiple Linear Regression Analysis examining the effect of potential confounding factors as well as Intervention Group on Intention to extract third molars (TMs)

						95% Confidence Interval for B	
Independent Variable	B [1]	SE [2]	Beta [3]	t	p	Lower	Upper
Intervention Group	-1.05	.44	-.25	-2.38	.020	-1.93	-0.17
Gender	-0.02	.51	-.01	-0.05	.959	-1.05	0.99
Age	0.11	.07	.44	1.55	.127	-0.03	0.24
Years qualified	-0.53	.50	-.31	-1.06	.293	-1.52	0.47
Post-Grad. Qualification	0.20	.56	.04	0.36	.718	-0.92	1.32
Courses attended last year	-0.02	.02	-.12	-1.04	.300	-0.06	0.02
TM Patients seen last year	-0.02	.03	.19	0.85	.400	-0.03	0.07
TM Patients seen last month	-0.47	.25	-.49	-1.90	.062	-0.96	0.02
TM Extractions last year	-0.08	.05	.50	1.72	.090	-0.01	0.17
TM Extractions last month	0.40	.34	.34	1.17	.246	-0.28	1.08

Dependent Variable: Intention to extract third molars; Method: Enter;

Adjusted $R^2 = 0.35$; $F(10,62) = 4.93$, $p<.001$

[1] Unstandardized Coefficients; [2] Standard Error; [3] Standardized Coefficients

Appendix II

Invitation Letter

Letter 1

Date Ext:

Dear

Effective practice?: a randomised trial of dissemination and implementation strategies for guidelines for the appropriate extraction of third molar teeth.

We are seeking the help of a selection of practices to assist us in the above study which is a collaboration between a number of Scottish Universities and the Scottish Council for Postgraduate Medical and Dental Education.

It will be of great help to us if your practice shows an interest in participating, as your contribution to the project will be of immense practical value.

Details of the study and what would be involved are given in the accompanying letter *(General Dental Practitioner Information Sheet)*. You can participate as an individual but it would be of great value if other members of the practice also became involved.

If you are interested in participating could you kindly complete the reply slip and return it in the envelope provided. If you require further information please do not hesitate to contact us on 01382 425754.

Yours sincerely

General Dental Practitioner Information Sheet

A Trial of Clinical Guidelines for the Appropriate Extraction of Third Molar Teeth

Introduction

The removal of third molar teeth is one of the most common surgical procedures. However many dentists are unclear as to which patients require extraction and which can be managed more conservatively. In order to address this area of uncertainty the Scottish Intercollegiate Guidelines Network (SIGN) are developing clinical guidelines to assist decision making regarding the correct management of third molars. All dentists in Scotland will receive copies of these during this year.

We would like to examine different methods of publishing guidelines and different ways of helping dentists remember and use the guidelines.

What will I have to do if I take part?

We would ask your help in identifying all patients aged between 16-24 years, who have attended your practice during a period 4 months **before** guideline publication. These records will then be searched by one of our research team in order to assess any patients with third molars problems, in order to record how they have been cared for. Following guideline publication by SIGN, the participating practices (approximately 60) will receive copies of the guidelines and the opportunity to attend lectures on the guidelines. Some dentists will be asked to incorporate structured forms to place in the patients' notes to help remind them of the guidelines, asked to join audit groups, or receive computer assisted learning (we will lend computers to this group if necessary). This will then be followed by a **second search** of patient records.

Participating dentists may also be interviewed by one of our researchers in order to identify reasons why one method is more or less successful than another at changing practice.

What are the risks of taking part?

All dentists in Scotland will receive copies of the guidelines and although these may represent best practice, guidelines are not prescriptive and the choice to follow, or not to follow, them in any individual case is yours. Therefore we do not see any risks in your participation.

We are interested in which method of providing the guidelines is most effective in influencing dentists' treatment decisions as a whole. **We are not interested in reporting individual cases and individuals (dentists or patients) will not be identified.**

We recognise that the record searches will cause inconvenience to your practice and we will request the help of your staff with this.

Are there any possible benefits?

It is hoped that adherence to the guidelines will improve the care provided. The information we obtain from the study will help us improve knowledge of the best methods to help dentists modify their practice appropriately and therefore improve the care patients receive.

Do I have to take part?

No, taking part is voluntary. If you do not wish to take part you do not have to give a reason. If you take part but later change your mind you can withdraw at anytime.

What do I do now?

If you wish to know more or have any questions please contact Chris Deery on 01382 425754 who will be happy to discuss the study with you.

Whatever decision you reach about participating we would be grateful if you could complete the enclosed form and return it to us.

Thank you very much for considering taking part in this research. Please feel free to discuss this information with any members of your practice team.

The Project Team
Dental Health Services Research Unit
Dundee Dental Hospital and School
Park Place
Dundee DD1 4HR

A: General Dental Practitioners Information Form

A Trial of Clinical Guidelines for the Appropriate Extraction of Third Molar Teeth

Personal Information

Name:	----------
Date of Birth:	----------
Graduating Institution:	----------
Year of Graduation:	----------

Practice Information

Address:	----------

Postcode:	----------
Telephone:	----------
Fax Number:	----------
E-Mail Address:	----------
No. of Dentists	----------
No. of Hygienists	----------

Signature: ----------

Date: ----------

B: General Dental Practitioner Consent Form

A Trial of Clinical Guidelines for the Appropriate Extraction of Third Molar Teeth

Name: .....

Practice Address: .....

.....

.....

.....

Telephone: .....

Please check above details and change if necessary

Please indicate your willingness to participate in this research project, by ticking the appropriate box below:

I am willing to participate in this research project ☐

I am not willing to participate in this research project ☐

I would like more information about the research project Please contact me ☐

Signature: .....

Date: .....

Please return to (a prepaid envelope is provided):

Dental Health Services Research Unit
Dundee Dental Hospital and School
Park Place
Dundee DD1 4HR

Appendix III

Course Participant Consent Form

Third Molar Project

Courses

Name: .....

Practice Address: .....

.....

.....

.....

Telephone: .....

Please check above details and change if necessary

Please indicate which course you are able to attend:

Glasgow	Thursday 8 June 2000	☐
Dundee	Thursday 15 June 2000	☐
Aberdeen	Thursday 22 June 2000	☐

Signature: .....

Date: .....

Please return by 15 April 2000 to:

Dental Health Services Research Unit
Dundee Dental Hospital and School
Park Place
Dundee DD1 4HR

Practice ID No:

Course Timetable

Third Molar Project
Glasgow Course

Venue: Wolfson Room
Postgraduate Centre
Glasgow Dental Hospital
378 Sauchiehall Street
Glasgow
G2 3JZ

1300 - 1400		Lunch
1400 - 1420	Dr Chris Deery (Programme Methodologist)	**Introduction** Guidelines - why SIGN up
1420 - 1445	Dr Maryam Bahrami (Clinical Research Fellow)	**Presentation** 'What is it all about and where are we?"
1445 - 1500		Coffee
1500 - 1600	Dr Liz Conner (Consultant in Oral Radiology)	**Presentation** "Radiographic assessment of third molar teeth"
1600 - 1700	Professor Graham Ogden (Consultant Oral & Maxillofacial Surgeon)	**Presentation** "How guideline influence the treatment decision"

Course Evaluation and Feedback Results

Q1: Responses to whether the course objectives were met or not, were 100% positive.

The objectives of this course were:

a) To understand the aims and objectives of third molar SIGN guideline
b) To promote best practice by using SIGN guideline to influence the treatment decision for unerupted and impacted third molar teeth

Q2: Rating for the course *i.e.* overall content, presentations, course organisation or design, relevance to general practice

Rating scales	Overall Content N=24 n (%)	Presentations N=24 n (%)	Course organisation / design N=24 n (%)	Relevance to general practice N=24 n (%)
1 = Outstanding	0	2(8)	0	0
2= Very good: fulfils educational objectives	21(88)	20(84)	21(88)	19(79)
3= More strengths than weakness	3(12)	2(8)	3(12)	4(17)
4= Significant weaknesses: some educational objectives not being met	0	0	0	0
5= poor. major revision of formal / content required	0	0	0	1(4)

Q4: Whether the course length was appropriate?

Appropriateness of the course length	
Rating scales	**Course length N=24 n(%)**
0= Not answered	1(4)
1= too long	1(4)
2= too short	1(4)
3= about right	21(88)

Q5: Whether the summary handouts were useful?

Usefulness of summary handouts	
Rating scales	**N=24 n(%)**
0= Not answered	1(4)
1= none given	15(63)
2= useful	5(21)
3= not useful	1(4)
4= None needed	2(8)

Q6: Whether the participant will recommend the course to a colleague?

Recommend to a colleague	
Rating scale	**N=24 n(%)**
0= Not answered	4(17)
1= Yes	17(71)
2= No	3(12)

Appendix IV

Proposed Audit Projects

Each group was offered two different project titles. They had the option of choosing one of these projects titles or they selecting any topic relating to third molars.

Group A

1. Are radiographs taken routinely for the management of third molar teeth? What types of radiographs are taken and what is the quality of the radiographs?

2. A retrospective audit, looking at the outcome of the third molar teeth which have been managed conservatively in the past and the outcome of this management.

Group B

1. Audit of quality of clinical assessment of unerupted and impacted third molar teeth.

2. What influences the decision as to where to refer patients with third molar problems?

Group C

1. What is the quality of record keeping relating to unerupted and impacted third molar teeth? Is the adequate and relevant information recorded?

2. If patients have a problem with one third molar tooth are the other present third molars examined?

Group D

1. Do referral letters provide adequate information on presentation, medical history and etc?

2. Are patients provided with sufficient information about their care including treatment they require, its delivery and what to expect.

Group E

1. How is pericoronitis managed? What signs/symptoms influence the treatment decision?

2. Do preoperative radiographs get sent to the operating surgeon on referral, and do they get returned to the original clinician at completion of treatment?

The selected audit projects by General Practitioners

Group 2A
Management of unerupted and impacted third molar teeth

Group 2B
An audit of the quality of clinical assessment in case notes of four GDPs of impacted and unimpacted third molars

Group 2C
Audit of charting and record keeping information with regard to third molar teeth in adult patients.

Group 2D
Are patients provided with sufficient information about their care including treatment they require its delivery, and what to expect in relation to wisdom teeth.

Group 2E
No audit

Group 4A
An audit of notes being taken on presence and prognosis of third molars in 18 to 29 years olds

Group 4B
Quality of assessment of lower third molars

Group 4C
No audit

Group 4D
A study to assess in patients aged 18- 27, their perception of care relating to the, management of third molars in general dental practice

Group 4E
No Audit

Appendix V

C: Record Form

A Trial of Clinical Guidelines for the Appropriate Extraction of Third Molar Teeth

Please record the names and dates of birth of patients whose date of birth lies between 1 January 1983 and 1 January 1975 who have attended your practice from<INSERT DATE>.

Name	Date of Birth	Name	Date of Birth

Letter of Visitation

Date

Dear

Effective practice?: a randomised trial of dissemination and implementation strategies for guidelines for the appropriate extraction of third molar teeth.

I would like to inform you of the date that we have allocated to visit your practice. If this date is not suitable we would appreciate if you could contact us as soon as possible so another one can be arranged. If we do not hear anything we will presume the visit can go ahead.

Your visit has been arranged for <INSERT DATE>

Please do not hesitate to contact me if there are any problems.

Yours sincerely

Appendix VI

Third Molar Data Entry Form

Date:

Practice ID No:

Dentist ID No:

Patient ID No: *CHI:*

Patient Name:

Patient Information

1	*Date of Birth*	☐☐☐☐☐☐
2	*Gender*	*Male* ☐ *Female* ☐
3	*Patient's postcode*	☐☐☐☐☐☐☐☐
4	*Date which the last medical history has been taken?*	☐☐☐☐☐☐
5	*Medical history*	*Heart* *Respiratory* *Blood disorder* *Diabetes* *Endocrine disorder* *Radiotherapy* *Allergy* *Hepatitis* *Other*
6	*Dental attendance*	*Regular attender* *Asymptomatic attender* *Symptomatic attender*
7	*Reason for the patient's attendance?*	*Check-up* *Toothache* *Problem with wisdom tooth/teeth* *Periodontal* *Fat face* *Atypical pain* *Orthodontic* *Treatment visit* *Other*

Third Molar Section

8	*Who first noticed the problem with the wisdom tooth/teeth?*	*Patient* *Dentist* *GMP*		
9	*Has the problem been recurrent?*		*Yes* ☐	*No* ☐
10	*Has the patient previously attended the dentist with the problem of wisdom tooth/teeth?*		*Yes* ☐	*No* ☐
11	*Have they had other wisdom teeth removed?*		*Yes* ☐	*No* ☐
12	*Duration of the last problem with the wisdom tooth/teeth?*			

Social History

13	*Does patient spend an extended period in a location where dentist is unavailable?*		*Yes* ☐	*No* ☐
14	*Smoke (number of cigarettes per day)*	*None* *2-10 per day* *Over 10 per day* *Not recorded*		
15	*Alcohol consumption*	*None* *Moderate* *High*		

Examination

16	*Extra-oral examination*	*Facial swelling* *Lymphadenopathy* *Trismus* *Pyrexic*				
17	*Third molar present*	18	*F* ☐	*P* ☐	28	*F* ☐ *P* ☐
		48	*F* ☐	*P* ☐	38	*F* ☐ *P* ☐

18		
18	Integrity of third molar tooth	Carious
		Sound
		Restored
19	Functional status	Functional
		Non-functional
20	Integrity of related 2nd molar	Carious/large restoration
		Sound

Radiographs

21	Radiographs	OPT
		Intra-oral
		Lateral oblique
		Not available
22	Standard of radiographs	Good
		Fair
		Poor
23	Angular position	MA
		DA
		V
		H
		inverted, transverse or Heterotopic
24	Impaction status:	partially covered by soft tissue
		completely covered by soft tissue
		partially covered by bone
		completely covered by bone
25	Shape and Curvature of the roots are	favourable
		not favourable
26	Stage of development	Crown only
		Crown +⅓ of root
		Crown +½ of root
		Fully developed root
27	Relation to 2nd molar	Class A (crown to crown position)
		Class B (crown to cervix position)
		Class C (crown to root position)

28	Retromolar space (38 & 48 only)	Class I retromolar space
		Class II space
		Class III ramus to body of mandible
29	Relation to IDC (38 & 48 only)	Deflection or diversion of IDC
		Narrowing of radiolucency of canal as nerve crosses the root
		Interruption of white line of the canal
30	Any history of problem with TMJD?	Yes ☐ No ☐
31	Clinical pictures (more than 1 poss)	Pericoronitis
		Cellulitis
		Caries in 2nd or 3rd molar
		Periapical abscess formation related to 3rd molar
		Untreatable pulpal/periapical pathology
		Periodontal disease involving impacted tooth
		Dentigerous cyst formation or other pathology
		External resorption of 3rd molar or related 2nd molar
		Atypical pain
		Orthodontic / imbrication of lower incisors
		Orthognahic surgery
		Atrophic mandible where 3rd molar tooth is present, with risk of fracture
		Overeruption
		Other

32	Management	Observation
		Drainage
		System antibiotics Metronidazole Amoxycillin Erythromycin Other
		Chlorhexidine mouth rinses
		Local dressing and lavage
		Hot salt mouthwashes
		Operculectomy
		Restoration
		Extraction of opposing tooth
		Easing off the cusps of opposing tooth
		Extraction under LA
		Extracion under LA + sedation
		Referred to specialist's centres
		Other

33	Has the patient encountered any problem with offended tooth, while on a waiting list for specialist clinic appointment?	Yes		No

34	Interim measure:	Systematic antibiotic
		Local dressing and lavage
		Chlorhexidine mouth rinses
		Drainage
		Operculectomy
		Easing off the cusps of opposing tooth
		Extraction of opposing tooth
		Other

28

18	*Integrity of third molar tooth*	*Carious*
		Sound
		Restored
19	*Functional status*	*Functional*
		Non-functional
20	*Integrity of related 2nd molar*	*Carious/large restoration*
		Sound

Radiographs

21	*Radiographs*	*OPT*
		Intra-oral
		Lateral oblique
		Not available
22	*Standard of radiographs*	*Good*
		Fair
		Poor
23	*Angular position*	*MA*
		DA
		V
		H
		inverted, transverse or Heterotopic
24	*Impaction status:*	*partially covered by soft tissue*
		completely covered by soft tissue
		partially covered by bone
		completely covered by bone
25	*Shape and Curvature of the roots are*	*favourable*
		not favourable
26	*Stage of development*	*Crown only*
		Crown +⅓ of root
		Crown +½ of root
		Fully developed root
27	*Relation to 2nd molar*	*Class A (crown to crown position)*
		Class B (crown to cervix position)
		Class C (crown to root position)

28	Retromolar space (38 & 48 only)	Class I retromolar space
		Class II space
		Class III ramus to body of mandible
29	Relation to IDC (38 & 48 only)	Deflection or diversion of IDC
		Narrowing of radiolucency of canal as nerve crosses the root
		Interruption of white line of the canal
30	Any history of problem with TMJD?	Yes ☐ No ☐
31	Clinical pictures (more than 1 poss)	Pericoronitis
		Cellulitis
		Caries in 2nd or 3rd molar
		Periapical abscess formation related to 3rd molar
		Untreatable pulpal/periapical pathology
		Periodontal disease involving impacted tooth
		Dentigerous cyst formation or other pathology
		External resorption of 3rd molar or related 2nd molar
		Atypical pain
		Orthodontic / imbrication of lower incisors
		Orthognahic surgery
		Atrophic mandible where 3rd molar tooth is present, with risk of fracture
		Overeruption
		Other

32	Management	Observation
		Drainage
		System antibiotics Metronidazole Amoxycillin Erythromycin Other
		Chlorhexidine mouth rinses
		Local dressing and lavage
		Hot salt mouthwashes
		Operculectomy
		Restoration
		Extraction of opposing tooth
		Easing off the cusps of opposing tooth
		Extraction under LA
		Extraction under LA + sedation
		Referred to specialist's centres
		Other

33	Has the patient encountered any problem with offended tooth, while on a waiting list for specialist clinic appointment?	Yes ☐	No ☐

34	Interim measure:	Systematic antibiotic
		Local dressing and lavage
		Chlorhexidine mouth rinses
		Drainage
		Operculectomy
		Easing off the cusps of opposing tooth
		Extraction of opposing tooth
		Other

38

18	*Integrity of third molar tooth*	*Carious*
		Sound
		Restored
19	*Functional status*	*Functional*
		Non-functional
20	*Integrity of related 2nd molar*	*Carious/large restoration*
		Sound

Radiographs

21	*Radiographs*	*OPT*
		Intra-oral
		Lateral oblique
		Not available
22	*Standard of radiographs*	*Good*
		Fair
		Poor
23	*Angular position*	*MA*
		DA
		V
		H
		inverted, transverse or Heterotopic
24	*Impaction status:*	*partially covered by soft tissue*
		completely covered by soft tissue
		partially covered by bone
		completely covered by bone
25	*Shape and Curvature of the roots are*	*favourable*
		not favourable
26	*Stage of development*	*Crown only*
		Crown +⅓ of root
		Crown +½ of root
		Fully developed root
27	*Relation to 2nd molar*	*Class A (crown to crown position)*
		Class B (crown to cervix position)
		Class C (crown to root position)

28	Retromolar space (38 & 48 only)	Class I retromolar space
		Class II space
		Class III ramus to body of mandible
29	Relation to IDC (38 & 48 only)	Deflection or diversion of IDC
		Narrowing of radiolucency of canal as nerve crosses the root
		Interruption of white line of the canal
30	Any history of problem with TMJD?	Yes ☐ No ☐
31	Clinical pictures (more than 1 poss)	Pericoronitis
		Cellulitis
		Caries in 2nd or 3rd molar
		Periapical abscess formation related to 3rd molar
		Untreatable pulpal/periapical pathology
		Periodontal disease involving impacted tooth
		Dentigerous cyst formation or other pathology
		External resorption of 3rd molar or related 2nd molar
		Atypical pain
		Orthodontic / imbrication of lower incisors
		Orthognahic surgery
		Atrophic mandible where 3rd molar tooth is present, with risk of fracture
		Overeruption
		Other

32	Management	Observation
		Drainage
		System antibiotics Metronidazole Amoxycillin Erythromycin Other
		Chlorhexidine mouth rinses
		Local dressing and lavage
		Hot salt mouthwashes
		Operculectomy
		Restoration
		Extraction of opposing tooth
		Easing off the cusps of opposing tooth
		Extraction under LA
		Extracion under LA + sedation
		Referred to specialist's centres
		Other

33	Has the patient encountered any problem with offended tooth, while on a waiting list for specialist clinic appointment?	Yes ☐	No ☐

34	Interim measure:	Systematic antibiotic
		Local dressing and lavage
		Chlorhexidine mouth rinses
		Drainage
		Operculectomy
		Easing off the cusps of opposing tooth
		Extraction of opposing tooth
		Other

48

18	*Integrity of third molar tooth*	*Carious*
		Sound
		Restored
19	*Functional status*	*Functional*
		Non-functional
20	*Integrity of related 2nd molar*	*Carious/large restoration*
		Sound

Radiographs

21	*Radiographs*	*OPT*
		Intra-oral
		Lateral oblique
		Not available
22	*Standard of radiographs*	*Good*
		Fair
		Poor
23	*Angular position*	*MA*
		DA
		V
		H
		inverted, transverse or Heterotopic
24	*Impaction status:*	*partially covered by soft tissue*
		completely covered by soft tissue
		partially covered by bone
		completely covered by bone
25	*Shape and Curvature of the roots are*	*favourable*
		not favourable
26	*Stage of development*	*Crown only*
		Crown +⅓ of root
		Crown +½ of root
		Fully developed root
27	*Relation to 2nd molar*	*Class A (crown to crown position)*
		Class B (crown to cervix position)
		Class C (crown to root position)

28	Retromolar space (38 & 48 only)	Class I retromolar space
		Class II space
		Class III ramus to body of mandible
29	Relation to IDC (38 & 48 only)	Deflection or diversion of IDC
		Narrowing of radiolucency of canal as nerve crosses the root
		Interruption of white line of the canal
30	Any history of problem with TMJD?	Yes ☐ No ☐
31	Clinical pictures (more than 1 poss)	Pericoronitis
		Cellulitis
		Caries in 2nd or 3rd molar
		Periapical abscess formation related to 3rd molar
		Untreatable pulpal/periapical pathology
		Periodontal disease involving impacted tooth
		Dentigerous cyst formation or other pathology
		External resorption of 3rd molar or related 2nd molar
		Atypical pain
		Orthodontic / imbrication of lower incisors
		Orthognahic surgery
		Atrophic mandible where 3rd molar tooth is present, with risk of fracture
		Overeruption
		Other

32	Management	Observation
		Drainage
		System antibiotics Metronidazole Amoxycillin Erythromycin Other
		Chlorhexidine mouth rinses
		Local dressing and lavage
		Hot salt mouthwashes
		Operculectomy
		Restoration
		Extraction of opposing tooth
		Easing off the cusps of opposing tooth
		Extraction under LA
		Extracion under LA + sedation
		Referred to specialist's centres
		Other
33	Has the patient encountered any problem with offended tooth, while on a waiting list for specialist clinic appointment?	Yes ☐ No ☐
34	Interim measure:	Systematic antibiotic
		Local dressing and lavage
		Chlorhexidine mouth rinses
		Drainage
		Operculectomy
		Easing off the cusps of opposing tooth
		Extraction of opposing tooth
		Other
35	Reviewed again by GDP?	Keep under observation
		Review again
		Refer for treatment
		Other

36	Seen by specialists:	Restorative
		Orthodontic
		Oral surgery
37	Specialist's treatment suggested include:	Extraction / surgical removal under LA
		Extraction / surgical removal under LA + sedation
		Extraction / surgical removal under GA
		Restoration
		Orthodontic treatment
		Observation

Appendix VII

Research Paper Accepted for Publication in British Dental Journal (Acceptance Date: October 2003)

Effectiveness of strategies to disseminate and implement clinical guidelines for the management of impacted and unerupted third molars in primary dental care, a cluster randomised controlled trial.

Bahrami M[1], Deery C[2], Clarkson JE[3], Pitts NB[4], Johnston M[5], Ricketts I[6], MacLennan G[7], Nugent ZJ[8], Tilley C[9], Bonetti D[10], Ramsay C[11].

[1]Clinical Research Fellow, Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee, DD1 4HR.

[2] Consultant in Paediatric Dentistry, Department of Paediatric Dentistry, Edinburgh Dental Institute, Edinburgh.

[3] Senior Lecturer Dental Primary Care/ Honorary Consultant; Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee, DD1 4HR.

[4]Professor, Director of Unit, Dental Health Services Research Unit; Dundee Dental Hospital; Park Place, Dundee, DD1 4HR.

[5] Professor in Psychology, School of Psychology, University of St Andrews, KY16 9JU.

[6] Professor of Assistive Systems & Healthcare Computing; Department of Applied Computing, University of Dundee.

[7]Statistician, Health Services Research Unit; University of Aberdeen.

[8]Statistician, Dental Health Services Research Unit, Dundee Dental Hospital: Park Place, Dundee, DD1 4HR.

[9]Research Fellow in Health Economics, Dental Health Services Research Unit. Dundee Dental Hospital; Park Place, Dundee, DD1 4HR.

[10]Research Fellow, School of Psychology, University of St Andrews, KY16 9JU.

[11]Research Assistant, Department of Applied Computing, University of Dundee.

Corresponding author:

Chris Deery
Dept of Paediatric Dentistry
Edinburgh Dental Institute
Lauriston Place
Edinburgh
EH3 9HA

Tel: 44 (0)131 536 4994

E-mail: chris.deery@lpct.scot.nhs.uk

Key words

Evidence based dentistry, guidelines, third molars, audit, computer aided learning, health economics.

Abstract

Objective: To investigate the effectiveness and cost-effectiveness of different guideline implementation strategies, using the SIGN Guideline 42 "Management of unerupted and impacted third molar teeth" (published 2000) as a model.

Design: A pragmatic, cluster RCT (2x2 factorial design).

Subjects: 63 dental practices across Scotland.

Clinical records of all 16-24 years old patients over two four month periods in 1999 (pre-intervention) and 2000 (post-intervention) were searched by a clinical researcher blind to the intervention group. Data were also gathered on the costs of the interventions.

Interventions: Group 1 received a copy of SIGN 42 Guideline and had an opportunity to attend a postgraduate education course. In addition to this, Group 2 received Audit & Feedback (A&F). Group 3 received a computer aided learning (CAL) package. Group 4 received A&F and CAL.

Principal Outcome Measurement: The proportion of patients whose treatment complied with the guideline.

Results: The weighted t-test for A&F versus no A&F (P=0.62) and CAL versus no CAL (P=0.76) were not statistically significant. Given the effectiveness results (no difference) the cost effectiveness calculation became a cost-minimisation calculation. The minimum cost intervention in the trial consisted of providing GDPs with guidelines and the option of attending PGCE courses. Routine data which subsequently became available showed a Scotland-wide fall in extractions prior to data collection.

Conclusion: In an environment in which pre-intervention compliance was unexpectedly high, neither CAL nor A&F increased the dentists' compliance with the SIGN Guideline compared to mailing of the guideline and the opportunity to attend a postgraduate course. The cost of the CAL arm of the trial was greater than the A&F arm. Further work is required to understand dental professionals' behaviour in response to guideline implementation strategies.

Background

It is being increasingly recognised that clinical care should be based on the best available evidence[1-2]. Unfortunately it has been estimated that only 15% of all clinical practice is based on sound research[3]. Part of this problem is a well recognised gap between the publication of research evidence and the adoption of this evidence in clinical practice[4]. The delay in uptake of a proven technology can be over a decade. The reasons for this delay are multi-factorial and suggested reasons include inertia, information overload, and difficulty in interpreting contradictory messages[5]. McGlone et al identified a number of barriers to change in dentistry: the profession's perception of patients (structural), patient's attitudes to dental health and the cost of dental care, fear of medico-legal action, treatment-funding system, time lost to practice when attending courses[5].

This problem in identifying and adopting evidence-based practice is as much a concern for health service planners and policy makers as it is for clinicians. Therefore together with the current emphasis on recognising the evidence, there is a need to promote best practice by changing clinician's behaviour in line with the evidence identified [5].

One method of presenting evidence in an accessible format is clinical guidelines. Clinical guidelines are defined as "systematically developed statements which assist in decision making about appropriate health care for specific clinical conditions" [6]. There are reports in the medical literature that clinical guidelines can improve the quality of care[7-8]. However passive dissemination of published guidelines alone is rarely effective in changing the clinical behaviour of practitioners, as many factors influence health

professionals' behaviour such as, organisational structure, peer group pressure and individual variation [9-10]. There is therefore a need to find effective implementation strategies to optimise the integration of research findings into current practice. Methods of disseminating and implementing research evidence which have been assessed in medical practice include continuing medical education, opinion leaders, audit and feedback, educational outreach, reminders and multi-faceted interventions [8-12]. However, it has been recognised that such strategies are not effective under all circumstances [7-13] and few studies have investigated their effectiveness in dentistry [5, 14-16].

Trials in dentistry have tended to look at referral practices. One trial evaluated the effectiveness of orthodontic referral guidelines and found that these did not influence the patient referral behaviour of general practitioners [14]. Another trial assessed three differing referral strategies from primary to secondary care for the treatment of impacted third molars and found a clinical algorithm (flowchart), performed better than then current practice, with a neural network based computer programme (third molar decision support system) performing least well [15].

The Scottish Intercollegiate Guidelines Network (SIGN) developed and published a clinical guideline relating to the management of impacted and unerupted third molar teeth in April 2000 (SIGN 43) [17]. This was identified as an appropriate subject for guideline development for a number of reasons *i.e.* variation in practice[15], the procedure is associated with considerable morbidity [18-20], the surgical removal of third molar teeth

is a common surgical procedure and is therefore associated with significant expenditure [21, 22].

The aim of this study was to evaluate the effectiveness of different implementation strategies for evidence-based clinical guidelines using SIGN 42 as a model [17]. In addition, the study conducted an economic evaluation of the dissemination and implementation strategies and examined behavioural factors mediating the effect of the interventions.

Methods

Ethical approval was obtained from The Multicentre Research Ethics Committee for Scotland (MREC) and the relevant local research ethics committees.

Study design

The study was conducted in dental practices across Scotland. The study was a pragmatic, cluster randomised controlled trial employing a 2x2 factorial design.

Sample size calculation

It was estimated that 4-5 patients per practice would have extractions during each data collection period based on data from the Scottish Dental Practice Board [23]. A sample size of 60 practices collecting information on 240 patients was required to detect a 20% reduction in inappropriate extractions from 60% to 40% [24] assuming 80% power and a 5% significant level. Based on the published trials in medical practice, an intra-class correlation coefficient (ICC) of 0.1 was incorporated into the power calculation to

account for the unit of randomisation being dental practices rather than individual practitioners[25].

Five hundred and sixty five general dental practices were randomly selected from the Scottish Dental Practice Board list and invited to participate in the trial by mail. In addition 41 practices who had previously participated in a Dental Health Services Research Units (DHSRU) trial were approached [26]. Sixty-three practices agreed to participate with written consent. Subsequently, 12 practices withdrew from the trial before randomisation. Therefore 51 of the original 63 volunteer practices were randomised into 4 groups by a statistician independent of the research team through computer generation of a random number sequence. Figure 1 presents the study profile.

The intervention groups were as follows:

Group 1 received a copy of SIGN 42 Guideline and had an opportunity to attend a postgraduate education course. In addition to this, Group 2 participated in Audit & Feedback (A&F). Group 3 received a specifically developed computer aided learning (CAL) package. Group 4 participated in A&F and received the CAL package.

Dentists completed a questionnaire based on social cognition models before and after the interventions, the results of which will be reported separately [27]. An economic evaluation of the relative cost effectiveness of each of the interventions was also performed.

Description of Interventions

Mailing of guidelines and Postgraduate (PG) courses

In April 2000, all dental practitioners in Scotland including those who were participating in the study received a copy of the guideline direct from SIGN. All GDPs recruited into the trial were invited to attend a postgraduate course. The courses ran on identical lines to those run by NHS Education for Scotland (formerly SCPMDE) and a practitioner's expenditure was reimbursed, as it would have been for any other course. It was planned to hold the 3 courses at regional centres. Attendance was not obligatory, since this was a pragmatic trial and due to lack of interest from the contacted dentists, one of the courses was cancelled.

Computer aided learning with decision support (CAL)

The CAL intervention strategy consisted of a laptop computer based support tool, with the potential to assist dental practitioners (Groups 3 and 4) in deciding on the appropriate treatment of third molars. The software was based solely on the SIGN Guideline, the computer delivered this advice in a multimedia format. The package was developed specifically for the trial.

Audit and feedback (A&F)

The participants in this arm of the trial (Groups 2 & 4) were divided into 11 groups according to the proximity of their practices to each other. The exact nature of the A&F was decided within each audit group and supported by the researcher with help and advice from the Scottish Council Dental Audit Tutor.

In each A&F group, one member was selected as a facilitator based on their previous experience and knowledge in carrying out an audit. The audit projects were conducted independently of each other on different aspects of clinical practice relating to third molar teeth.

Outcome measurement

The outcome measure was the proportion of patients whose treatment complied with the guideline. This was assessed by 2 independent researchers and any disagreements were resolved by discussion.

Dental practitioners were asked to identify all 16-24 year old patients who attended their dental surgery over two 4-month periods in 1999 pre- and 2000 post- intervention. A Researcher blind to the intervention groups retrieved these records and transcribed the reason for attendance and treatment on a previously piloted form. Pre-intervention data were collected from 49 practices and post-intervention data from 47 practices.

Statistical analysis

The trial had a factorial design, which provided an opportunity to study and test for interactions between the interventions if they existed.

The level of compliance with the guideline between intervention groups was tested by means of a cluster level analysis using a weighted *t*-test [28]. To control for heterogeneity between patients (case-mix) and practices, multilevel regression analysis was also used. All analyses were on an intention to treat basis.

Results

Practice recruitment

Of 565 dentists invited to participate, at least one dentist from 63 practices volunteered to participate. Twelve of these practices withdrew before pre-intervention data gathering. The 51 remaining dental practices were randomised to four groups. During the period between pre-intervention and post-intervention data collection four practices withdrew from the trial, leaving 47 available for post-intervention data collection. The reasons given for withdrawal were too busy, not interested, moving practice, refurbishment of practice and a change of mind. There were no significant differences between the dentists who withdrew from study and those who continued, in terms of their age (t = -1.34, p=0.2); gender (χ^2 = 0.15, df=1, p = 0.7); postgraduate qualifications (χ^2 = 2.24, df=1, P = 0.1) and their intervention group (χ^2 = 4.24, df=1, P = 0.2).

Of the 51 pre-intervention practices, 47 provided data. Of the 47 post-intervention practices 46 had data relevant third molar teeth. The numbers of practices or clusters per group are shown in Figure 1. The characteristics of the participating dentists are reported in Table 1.

Of the dentists recruited, 23 (45%) attended one of the two post-graduate courses.

Pre-intervention, data were collected for 3342 (M=1885, F=1457) patients with a mean age of 21.7 years (SD 2.2, range 14-25.5 yrs) compared with 1935 (M= 880, F=1055) patients at the post-intervention stage with a mean age of 21.8 years (SD 2.1, range 16.6-25 yrs) (Table 3). The proportion of patients with a problem with their third molar

teeth was 7% before compared with 22% after intervention (P=0.0001). Female patients were significantly more likely to present with third molar problems than males in both the pre and post-intervention phases (P=0.0001) (Table 2).

Outcome assessment

Overall compliance with the guideline (Evidence-Based Practice) at pre-intervention stage was assessed to be 74% and this increased to 78% at post-intervention.

Comparison of pre-intervention data between the groups indicates that there was no apparent imbalance between groups at baseline (Table 3). The level of adherence to the guideline prior to any discussion, as assessed by the two independent examiners was 74% and 68% pre-intervention and 78% and 75% post-intervention, respectively.

The weighted *t*-test for A&F versus no A&F post-intervention was not statistically significant (p=0.62) neither was that between CAL versus no CAL (p=0.76).

From the multilevel analysis the odds ratio of compliance with guidelines for dentists who experienced A&F versus those who did not was 1.28 (95% CI 0.62 - 2.63) and this compared with an odds ratio of 0.84 (95% CI 0.88 - 1.74) for the CAL dentists versus no CAL. Neither difference was statistically significant (Table 4).

To account for case-mix "effect modifiers" in the multilevel model, pericoronitis, caries and pulpal pathology were included. All patients with pulpal pathology were treated in accordance with the guideline. Only one case presenting with caries was not treated in

accordance with the guideline. For patients presenting with pericoronitis there was an increase in compliance with guidelines but the effect was not different across groups.

The rate of third molar extractions decreased for 16-24 years olds after the introduction of the SIGN Guideline. This reduction was statistically significant between the pre- and post-intervention phase of this study (37% to 27%, p=0.02) and was consistent across the groups.

Concurrent with the study data, data from the NHS management information and dental accounting system (MIDAS) shows a concomitant reduction in both surgical and non-surgical wisdom tooth extraction rates in Scottish general dental practices during the experimental period (figs 2 and 3) [29].

Cost effectiveness

With no evidence of an effect, the planned cost effectiveness analysis became a cost-minimisation calculation. After controlling for scale effects, the driver behind cost differences between the groups was CAL. Sensitivity analysis ignoring PC purchase cost and delivery costs attenuated these differences. There were substantial differences in the costs of the interventions with the CAL arm of the trial costing £482.34 per dentists and the A&F per dentist £216.51. However, a sensitivity analysis ignoring PC purchase and delivery costs attenuated these differences

Discussion

The design of the study was a clustered randomised controlled trial with an intention to treat analysis. The design provided an opportunity to compare the relative effectiveness of different interventions and balance any modifying effects associated with the context, study population or outcomes [30].

The main finding of this trial is that neither A&F nor CAL was more effective than mailing and attendance at a postgraduate course, in increasing the general dental practitioners compliance with the SIGN guideline for management of impacted and unerupted third molar teeth. This finding is not unique in the primary care setting. A trial published by Eccles and colleagues suggested that A&F does not change behaviour for primary-care radiology referrals [31].

A previous evaluation of CAL, to develop clinical decision-making skills found no evidence of an effect on dentists' treatment decision-making behaviour[16].

An unexpected finding was that there was good adherence with the guideline recommendations at baseline (74%). This level of compliance is higher than would have been predicted from other research evaluating compliance in the UK with the USA's National Institutes of Health (NIH) criteria [25]. It may be that this and other publications have influenced knowledge levels and practice since 1996. There has, over recent years, been discussion of the appropriateness of the extraction of impacted third molars in the national and dental media. For example, an Effectiveness Matters publication from the Centre for Reviews and Dissemination at the University of York

was published in 1998 [21]. The findings of the review were based on other reviews and research findings that have been widely reported in the dental press [32-35] and this may have influenced Scottish dentists' behaviour (Figs 2 & 3). If publications relating to the extraction of third molars have influenced care it is one of the few examples of relatively passive dissemination and implementation being effective in altering clinical practice.

The lack of a difference between the interventions may be a true lack of effect or it may be a reflection of the high level of compliance at baseline which may have produced a "ceiling effect", where no greater improvement within the group was possible.

Another significant confounder may be that the volunteer dental practitioners who consented to participate in our trial may not be representative of the Scottish GDP population possibly having a particular interest in research or the subject of the research. They may be more likely to participate than others and may manage their patients more effectively, rendering the sample less representative of the target population and possibly introducing bias. Current recommendations for the conduct of randomised trials, require each participant be provided with enough information to allow for well-informed consent [36]. Accordingly, there can be a systematic tendency to recruit a biased group.

A second explanation for the high compliance could be the "Hawthorne Effect", a social placebo response. Practitioners may have acted differently because they were participating in the trial. Forty-two percent fewer patients were seen post intervention

compared with pre-intervention, with a greater proportion having a third molar problem. The authors do not have an explanation for this finding. An attempt was made to prevent the "Hawthorne Effect" by keeping the process of data collection as unobtrusive as possible [25]. Other explanations for this finding could be that the dentists participating in the trial were more confident in their management of the condition as a result of the guideline, or that there was improvement in case-note keeping as a consequence of their participation in trial.

Unfortunately, a number of practices withdrew from the study, particularly during the first phase. The reasons given for withdrawal from the study indicate that they were principally for reasons other than the demands of the project. The loss of practices seems not to have had a significant effect on the study's power to detect any difference between the interventions, if one was present. The power calculation used to estimate sample size in this study drew its estimates of probable ICC from comparable studies carried out in medical primary care. In practice, the ICC in the study was actually 0.15. This suggested that the study had reasonably high power to detect a 20% difference between the interventions if it was present. This can be used to inform future sample size calculations in studies of this type.

The results of the study examining the mediators of behavioural change demonstrated that dentists' cognition were demonstrably affected by the interventions with those who received the A&F intervention increasing their third molar-related knowledge significantly more than dentists who did not receive A&F [27]. However, the cognitions that were changed did not relate to extraction behaviour.

The dissemination and implementation of guidelines are extremely costly. We compared these very large costs with any benefits that accrued as a result of the interventions. As we were unable to detect any benefit from the interventions the most cost-effective choice is the least-cost option, in this case mailing and the opportunity to attend a postgraduate course. The CAL intervention was found to be the most expensive.

As recognised by McGlone *et al* our knowledge of altering professional practice in dentistry and particularly as it relates to the effective utilisation of clinical guidelines is somewhat limited. There is an urgent need for further research in the area of altering professional practice in dentistry, which will ensure efficient and effective use of limited resources, with the potential to improve the delivery of care by promoting best practice [5].

Conclusion

In an environment where pre-intervention compliance was unexpectedly high, there were no evidence that CAL and A&F increased general dental practitioners compliance with the SIGN guideline for the management of impacted and unerupted third molar teeth compared with mailing and the opportunity to attend a postgraduate course alone. Related research suggests that these interventions may act as reinforcement of the guideline messages [27].

Acknowledgements

Professor Graham Ogden, Dundee Dental Hospital and School, University of Dundee, for help with the development of the CAL Package and the postgraduate courses.

For additional statistical advice Marion Campbell of the Health Services Research Unit (HSRU), University of Aberdeen.

Professor Jeremy Grimshaw, Ottawa Health Research Institute, Canada (formerly HSRU) and Professor Trevor Sheldon, University of York for methodological advice.

Elaine Humphries, Audit Facilitator, NHS Education for Scotland her support.

The following General Dental Practitioners assisted with the development of the CAL package. Dr Iain Hunter, Dr Andrew Wright, Dr Jim McCafferty, Dr Colin Yule.

We would also like to thank everyone involved in the study, particularly Marilyn Laird, Jenny Johnson, Louise Cardno and the participating dental practices.

This project was supported by Scottish Executive's Chief Scientists Office and NHS R&D Programme (Trial R2-64). The Dental Health Services Research Unit is core funded by the Chief Scientist Office of the Scottish Executive Health Department. The views expressed are those of the authors.

References

1. NHS. A first class service: quality in the new NHS. England: NHS 1998.

2. Richards D, Lawrence A. Evidence-based dentistry. Br Dent J 1995; 179: 270-273.

3. Scottish Home and Health Department. Clinical guidelines. Edinburgh, UK, 1993.

4. Davidoff F, Haynes B, Sackett D, Smith R. Evidence-based medicine. Br Med J 1995; 310: 1085-1086.

5. McGlone P, Watt R, Sheiham A. Evidence-based dentistry: an overview of the challenges in changing professional practice. Br Dental J 2001: 190: 636-639.

6. Institute for Medicine Committee on Clinical Practice Guidelines, Guidelines for clinical practice: from development to use. National Academy Press: Washington DC, USA, 1992.

7. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: A systematic review of rigorous evaluations. Lancet 1993; 342: 1317-1322.

8. Effective Health Care Bulletin. Implementing clinical guidelines: can guidelines be used to improve clinical practice. York: University of York, 1994; 8, 1-12.

9. Lomas J. Words without action? The production, dissemination, and impact of consensus recommendations. Ann Rev Public Health. 1991; 12: 41-65.

10. Effective Health Care Bulletin; Getting evidence into practice. York: University of York,1999; 5: 1

11. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J. Using clinical guidelines. Brit Med J 1999; 318: 728-730.

12. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD. Thomson MA. Getting research finding into practice: closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. Br Med J 1998; 317: 465-468.

13. Oxman A. No magic bullets: a systematic review of 102 trials of interventions to help health care professionals deliver services more effectively or efficiently. Can Med Associ J 1995; 153: 1423-31.

14. O'Brien K, Wright J, Conboy F, *et al.* The effect of orthodontic referral guidelines: a randomised controlled trial. Br Dent J 2000; 188: 392-397.

15. Goodey RD, Brickley MR, Hill CM, Shepherd JP. A controlled trial of three referral methods for patients with third molars. Br Dent J 2000; 189: 556-560.

16. Kay EJ, Silkstone B, Worthington HV. Evaluation of computer aided learning in developing clinical decision-making skills. Br Dent J 2001; 190: 554-557.

17. Scottish Intercollegiate Guidelines Network (SIGN). Management of unerupted and impacted third molar teeth, A National Clinical Guideline. SIGN 2000.

18. Carmichael FA, McGowan DA. Incidence of nerve damage following third molar removal: A West of Scotland Oral Surgery Research Group Study. British Journal of Oral and Maxillofac Surg 1992; 30: 78-82.

19. Blackburn CW, Bramley PA. Lingual nerve damage associated with the removal of lower third molars. Br Dent J 1989; 167:103-106.

20. Chiapasco M, De Cicco L, Marrone G. Side effects and complications associated with third molar surgery. Oral Surg Oral Med Oral Pathol 1993; 76: 412-420.

21. Effectiveness Matters: Prophylactic removal of impacted third molars: is it Justified? The Centre for Reviews and Dissemination, University of York, 1998; 3: issue 2.

22. Sadler A, Davidson M, Houpis C, Watt-smith S, Specialist practice for minor oral surgery: a comparative audit of third molar surgery. Br Dent J 1993; 174:273-277.

23. Dental Practice Board Annual Report. Scottish Dental Practice Board. Trinity Park House, Edinburgh, 1997.

24. Brickley MR, Shepherd JP. An investigation of the rationality of lower third molar removal, based on USA National Institutes of Health criteria. Br Dent J 1996; 180: 249-254.

25. Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. Family prac 2000: 17 (suppl): 1-16.

26. Deery C, Fyffe HE, Nuttall NM, Nugent ZJ, Pitts NB. The dental caries status of Scottish adolescents reported to be regular attenders. Initial results from a primary dental care based research network. Br Dent J 1999; 187: 95-99.

27. Bonetti D, Johnston M, Pitts N, Deery C, Ricketts I, Bahrami M. Ramsay C. Johnston J. Can psychological models bridge the gap between clinical guidelines and clinicians' behaviour? A randomised controlled trial of an intervention to influence dentists' intention to implement evidence-based practice Br Dent J. 195.7: 403-407.

28. Bland J, Kerry S. Statistical Notes: Weighted comparison of means. Br Med J 1998: 316: 129.

29. NHS Management Information and Dental Accounting System (MIDAS). Practitioner Services Division, Edinburgh.

30. Grimshaw JM, Freemantle N, Langhorne P, Song F. Complexity and systematic reviews. Report to the US Congress of Technology Assessment. Office of Technology Assessment, Washingon DC, 1995.

31. Eccles M, Steen N, Grimshaw J. Thomas L, McNamee P, Soutter J, Wilsdon J. Matowe L, Needham G, Gilbert F, *et al.* Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial. Lancet 2001: 357:1406-1409.

32. Lands DP. The relationship between dental health and variations in the level of third molar removals experienced by populations. Community Dent Health 1998; 15: 67-71.

33. Song F, Landes D, Glenny A-M, Sheldon TA. Prophylactic removal of impacted third molars: an assessment of published reviews. Br Dent J 1997; 182: 339-46.

34. Daley T. Third molar prophylactic extraction: a review and analysis of the literature. General Dentistry J 1996; 44: 310-320.

35. Mercier P, Precious D. Risks and benefits of removal of impacted third molars. Journal of Oral and Maxillof Surg 1992; 21: 17-27.

36. Edwards SJL, Lilford RJ, Braunholtz DA, *et al*. Ethics of randomised trials. In Black N, Braizier J, Fitzpatrick R, Reeves B: Health services research methods. A guide to best practice. pp 98-107, London: BMJ publication, 1998.

Table 1

Participants Characteristics					
No of Dentists = 51		Gender		Mean age	Year of qualification
		Male	Female		
Groups		41	10	42 (SD 7.8)	1965-1997 (Median1981)
Group 1	Non- intervention /control	8	4	44	1971-1989
Group 2	A&F	12	1	38	1966-1987
Group 3	CAL-DS	9	4	44	1968-1997
Group 4	A&F + CAL-DS	12	1	41	1965-1993

Table 2

The gender distribution of patients with and without third molar problems before and after intervention								
Patients	Pre- intervention (Overall no of patients) **N=3342**				Post- intervention (Overall no of patients) **N=1934**			
	Male n(%)	**Female n(%)**	**Total n(%)**	***P value**	**Male n(%)**	**Female n(%)**	**Total n(%)**	***P value**
with third molar problem	98 (40)	146 (60)	244 (100)	0.0001	161 (38)	265 (62)	426 (100)	0.0001
without third molar problem	1787 (58)	1311 (42)	3098 (100)		719 (48)	789 (52)	1508 (100)	

*Pearson Chi-square

Table 3

Weighted mean percentage compliance with guidelines for all groups				
Groups	**Pre**		**Post**	
	% (SD)	**(95% CI)**	**% (SD)**	**(95% CI)**
Control	77 (12)	(70 - 85)	81(18)	(70 - 92)
Audit & Feedback	77 (18)	(66 - 86)	78 (10)	(73 - 84)
CAL	70 (24)	(56 - 84)	73 (25)	(59 - 88)
Audit & Feedback plus CAL	75 (24)	(62 - 88)	82 (23)	(79 - 95)

Table 4

Percentage compliance with guidelines for dentists who experienced A&F versus those who did not and similarly for CAL			
	Post-intervention % (SD)	**Odds ratio (95% CI)**	**P-value**
No Audit & Feedback	77.2 (21.4)	1.28 (0.62 to 2.63)	0.51
Audit & Feedback	79.4 (15.1)		
No CAL	79.3 (13.2)	0.84 (0.88 to 1.74)	0.65
CAL	77.3 (23.7)		